ZAG-1/ZEB and EGL-44/TEAD form a negative feedback loop to safeguard the choice of cell fate

Chaogu Zheng, Felix Qiaochu Jin[1], Brian Loeber Trippe[1], and Martin Chalfie[2]

Department of Biological Sciences, Columbia University, New York, New York, 10027

[1]These authors contributed equally.

[2]Lead contact and to whom correspondence should be sent at:

Department of Biological Sciences

1012 Fairchild, MC#2446

Columbia University

1212 Amsterdam Avenue

New York, NY 10027

Phone: 212-854-8870

Fax: 212-865-8246

Email: mc21@columbia.edu.

Summary

Terminal differentiation generates the specialized structures and functions that allow postmitotic cells to acquire their distinguishing characteristics. This process is thought to be controlled by transcription factors called "terminal selectors" that directly activate a set of effector genes. Using the mechanosensory touch receptor neurons (TRNs) in *Caenorhabditis elegans*, we extend this concept by identifying non-selector regulators of cell fate, which promote TRN fate indirectly by safeguarding the stability and activity of the selectors. In particular, the ZEB family transcriptional repressor ZAG-1 promotes TRN fate by inhibiting the TEA domain transcription factor EGL-44 and its binding partner EGL-46. EGL-44 and EGL-46 inhibit the TRN fate and simultaneously induce the fate of the multidendritic nociceptor FLP neurons. Since TRN and FLP share the same terminal selectors and a common ground state, the mutual inhibition between ZAG-1 and EGL-44 forms a negative feedback loop to regulate the choice between the two neuronal fates.

Introduction

The terminally differentiated state or cell fate of neurons distinguishes them from other neurons through specialized features and functions and the expression of effector genes ("terminal differentiation genes") (Hobert, 2011). Although individual effector gene may not be expressed in a cell type-specific manner, the collective expression of a battery of terminal differentiation genes serve as a "signature" for the cell fate. For example, the cell fate of mammalian photoreceptor is defined by the expression of over 600 genes, including photoreceptor-specific retinol-binding proteins, rhodopsin, G protein Gnat1, phosphodiesterase,

and others that are directly involved in light detection (Hsiau, et al., 2007; Blackshaw, et al., 2001).

The expression of the terminal differentiation genes is thought to be activated by terminal selectors (mainly transcription factors) that act alone or in combination (Hobert, 2011). Examples of such selectors are the C2H2 zinc finger transcription factor CHE-1, which determines the fate of gustatory ASE neurons in *C. elegans* (Etchberger, et al., 2007), the homeodomain transcription factor Crx, which controls photoreceptor fate (Corbo, et al., 2010), and the C4 zinc finger transcription factor Nurr1 and the homeodomain protein Pitx-3, which cooperatively direct the differentiation of mouse midbrain dopaminergic neurons (Martinat, et al., 2006). Terminal selectors directly up-regulate the transcription of various terminal differentiation genes through common *cis*-regulatory elements (Hobert, 2016; Hobert, 2011); thus, removing the selector genes leads to the loss of the expression of differentiation genes and the loss of neuronal fates.

Two problems arise in considering this model for differentiation. First, not all transcription factors that determine cell fate, however, can be classified as "terminal selectors", because some may indirectly regulate terminal differentiation genes. Although the loss of such non-selector regulators also causes the absence of a particular cell fate, the mechanism underlying their action is not well understood. Second, the same terminal selectors are often expressed in several distinct types of neurons and are required for their differentiation into distinct fates, yet they only promote one cell fate in one particular neuron (Hobert, 2016). Such shared selectors presumably work in coordination with other fate-specific selectors. Although the overlapping expression pattern of many transcription factors could give rise to unique

combinatorial code for each cell fate, the mechanism that restricts the function of shared terminal selectors during the specification of particular fate is not clear.

To address these issues, we have studied the choice of cell fate that generates either the touch receptor neurons (TRN) or the FLP neurons in *C. elegans*. The six TRNs are mechanosensory neurons that detect gentle mechanical stimuli along the body (Chalfie and Sulston, 1981), and the two FLP neurons are multidendritic nociceptors that sense harsh touch, noxious temperature, and humidity (Russell, et al., 2014; Chatzigeorgiou and Schafer, 2011; Chatzigeorgiou, et al., 2010; Kaplan and Horvitz, 1993). The heterodimer of the POU homeodomain transcription factor UNC-86 and the LIM homeodomain transcription factor MEC-3 acts as a terminal selector that promotes TRN cell fate (Xue, et al., 1993). Specifically, the heterodimer activates a set of TRN terminal differentiation genes (the mechanosensory channel genes *mec-4* and *mec-10*, the tubulin genes *mec-7* and *mec-12*, the tubulin acetyltransferase gene *mec-17*, and others) by binding to conserved regulatory elements in their proximal promoters (Zhang, et al., 2002; Duggan, et al., 1998).

In addition to directing the differentiation of the TRNs, MEC-3 and UNC-86 are expressed in and needed for the differentiation of two pairs of multidendritic neurons, the embryonic and anterior FLP neurons and the postembryonic and posterior PVD neurons. The expression of *unc-86* and *mec*-3 in these cells does not lead to the normal expression of many TRN terminal differentiation genes, resulting in cells whose fate is very different from the TRNs (Finney and Ruvkun, 1990; Way and Chalfie, 1988). Our previous work suggested that the TEA domain transcription factor EGL-44 and the zinc-finger protein EGL-46 in FLP neurons prevented these cells from acquiring the TRN fate (Wu, et al., 2001). Thus, the active inhibition of TRN fate by EGL-44 and EGL-46 in FLP neurons distinguished these two types of neurons

that express common selector genes. However, whether EGL-44 and EGL-46 also activate a FLP-specific genetic program is unclear.

In this paper, we extend this model of TRN and FLP neuronal determination, by asking whether other transcription factors help promote the TRN fate and cause the cells to differentiate away from the FLP fate. We screened for the effects on TRN development of the loss of 392 transcription factors using RNA interference and identified LDB-1 and ZAG-1 as positive TRN fate regulators. The LIM domain-binding protein LDB-1 binds to MEC-3 protein and stabilizes it post-transcriptionally, thus contributing to the activation of both TRN and FLP fate. The Zn-finger homeodomain protein ZAG-1 is expressed in TRNs but not FLP neurons. ZAG-1 promotes the TRN fate not by directly activating the TRN terminal differentiation genes but by preventing *egl-44* and *egl-46* expression, which blocks the activation of TRN genes by UNC-86/MEC-3. In FLP neurons, EGL-44/EGL-46 complex simultaneously inhibits the TRN genes and activates FLP genes. EGL-44/EGL-46 also represses the expression of *zag-1* in FLP neurons, and this negative feedback loop establishes a bistable switch between TRN and FLP fates. Our work suggests that UNC-86/MEC-3 serves as a ground-state selector resulting in a common state in both TRNs and FLP neurons, and that individual fates are subsequently controlled by ZAG-1 and EGL-44/EGL-46 acting in a bistable switch. In particular, ZAG-1 regulates the choice of cell fate by inhibiting fate inhibitors and safeguarding the activation of terminal differentiation genes.

Results

An RNAi screen for transcription factors affecting TRN cell fate

The anterior ALML and ALMR neurons and the posterior PLML and PLMR neurons are two pairs of bilaterally symmetric TRNs generated in the three-fold embryo, i.e., late in

embryonic development. Two other TRNs, the AVM and PVM neurons, arise post-embryonically (Chalfie and Sulston, 1981; Sulston and Horvitz, 1977). To search systematically for transcription factors specifying TRN fate, we knocked down the expression of transcription factor genes using RNAi and looked for animals that no longer expressed *mec-17p::RFP*, a TRN marker (see STAR Methods for details). Using RNAi against *unc-86* as a positive control, we tested various genetic backgrounds that were previously found to enhance the effects of RNA interference and found that *eri-1; lin-15B* mutants had the highest penetrance for the loss of RFP expression. About 80% of these animals did not express RFP in either of the two ALM neurons when treated with *unc-86* RNAi (Figure 1A and B). Since the posterior TRNs (PLM neurons) were less affected by the RNAi treatment (Figure 1B), we focused on the disappearance of *mec-17p::RFP* expression in the ALM neurons in the screen.

Among the 443 bacterial clones expressing dsRNA against 392 transcription factors and associated proteins (Table S1), we identified 14 genes that were required for the expression of TRN markers (Figure 1). Four genes (*unc-86*, *mec-3*, *ldb-1*, and *ceh-20*) were previously known to affect the expression of TRN terminal differentiation genes. We examined null mutants for all the remaining ten genes but could only confirm the loss of *mec-17p::RFP* expression in *zag-1* mutants.

We were surprised that RNAi against nine genes affected TRN expression through six rounds of testing but loss-of-function mutations in these genes did not. These false positive results are unlikely to result from the specific genetic background of the RNAi strain, since mutation of several of the genes (*zip-4*, *hmbx-1*, and *nhr-119*) in *eri-1; lin-15B* animals did not affect TRN fate. Activating the RNAi pathway non-specifically by RNAi against GFP in those triple mutants (e.g. *zip-1; eri-1; lin-15B* mutants) did not cause the loss of *mec-17p::RFP*

expression either. Therefore, we reason that the discrepancy between the RNAi and mutant phenotypes was likely due to mistargeting of the dsRNAs.

The LIM domain-binding protein LDB-1 promotes TRN fate by stabilizing MEC-3

ldb-1 encodes the only *C. elegans* homolog of nuclear LIM domain-binding protein NLI/Ldb1/CLIM2 and binds to LIN-11 and MEC-3, two of the founding members of the LIM homeodomain domain protein family (Gupta, et al., 2003; Cassata, et al., 2000; Freyd, et al., 1990; Way and Chalfie, 1988). Cassata *et al.* (2000) found that RNAi against *ldb-1* eliminated the expression of TRN gene *mec-2* without affecting *mec-3* expression. We found that an *ldb-1* null mutation, which caused early larval arrest, eliminated the expression of several TRN markers (*mec-4p::GFP*, *mec-7p::GFP*, *mec-17p::RFP*, and *mec-18p::GFP*) in both ALM and PLM neurons in those arrested L1 animals (Figure 2A and B). The loss of *ldb-1* also eliminated the expression of FLP fate marker *sto-5p::GFP* in FLP neurons (Figure S1A). Thus, *ldb-1* is required for the specification of TRN and FLP fates, both of which depend on MEC-3 as a terminal selector. We found that the expression of *mec-3p::RFP* reporter persisted in *ldb-1* mutants (Figure 2C), but quantitative measurements of *mec-3* mRNA levels using single molecule fluorescent *in situ* hybridization (smFISH) revealed that *mec-3* transcription was moderately reduced in *ldb-1* mutants compared to the wild type animals (Figure 2D). Therefore, our data suggest that LDB-1 is not only required for the activation of MEC-3 target genes but is also needed for optimal *mec-3* expression. This result is consistent with the known *mec-3* autoregulatory activity (Topalidou, et al., 2011; Xue, et al., 1992).

The physical interaction between LDB-1 and MEC-3 was confirmed using a yeast two-hybrid assay (Figure S1B-D). This interaction presumably enables MEC-3 to activate downstream TRN differentiation genes in a manner similar to the function of the mouse LDB-1

homologs NLI and CLIMs, which potentiate the transactivation by LIM-homeodomain proteins, like Lhx3 and Isl1 (Bach, et al., 1997; Jurata, et al., 1996). Although LDB-1 is not needed for the DNA binding capacity of MEC-3 at least *in vitro* (Xue, et al., 1993), the activation of downstream genes by MEC-3 may depend on LDB-1 *in vivo*.

Interestingly, the binding of MEC-3 to UNC-86 was very weak compared to the MEC-3/LDB-1 interaction in our yeast two-hybrid system (Figure S1D), which was unexpected given earlier results of electrophoretic mobility shift assay (EMSA) that found stable association of UNC-86/MEC-3/DNA complex (Xue, et al., 1993). Our results suggest that the UNC-86/MEC-3 heterodimer is not stable without the target DNA. Formation of this heterodimer may mostly occur on the promoter of TRN genes, where the binding sites for UNC-86 and MEC-3 are adjacent.

We also found that the strong association of LDB-1 with MEC-3 is required for the protein stability of MEC-3; unlike the *mec-3* promoter reporter, the expression of *mec-3::GFP* translational fusions was strongly reduced in *ldb-1* mutants (Figure 2E). These results suggest that the MEC-3 protein is highly unstable in the absence of LDB-1. In fact, treatment with the proteasome inhibitor MG-132, which reduces the degradation of ubiquitin-conjugated proteins, significantly increased the level of MEC-3::GFP protein in *ldb-1* mutants (Figure 2F), supporting the idea that LDB-1 stabilizes MEC-3 and prevents its degradation. Similarly, the Drosophila LDB-1 homolog Chip binds to and stabilizes the LIM-homeodomain protein Apterous, and the dissociation of Chip from Apterous triggers its degradation (Weihe, et al., 2001). These results suggest a common function for the nuclear LIM interactor (Ldb/NLI/Clim) in regulating LIM-homeodomain protein homeostasis.

We confirmed the results of Cassata *et al.* (2000) that *ldb-1* is expressed in many neurons, vulval cells, and some muscle cells, but also found that it was expressed in the TRNs in a *mec-3*-dependent fashion (Figure 2G). This latter result suggests that MEC-3 positively regulates its own activity and stability by activating the expression of its binding partner. Thus, MEC-3 autoregulation occurs at both the post-translational and functional levels.

ZAG-1 is required for the expression of TRN fate markers independently of MEC-3

zag-1 encodes the sole *C. elegans* ZEB transcription factor. Human ZEB transcription factors, which like ZAG-1 have a homeodomain flanked by clusters of C2H2-type zinc fingers, induce the epithelial to mesenchymal transition (EMT) and are essential for normal embryonic development (Vandewalle, et al., 2009). Mutations in human ZEB genes cause defects in neural crest development and are linked to malignant tumor progression (Vandewalle, et al., 2009). *C. elegans zag-1* prevents the posterior post-embryonic TRN neuron PVM from adopting the morphology of the multidendritic nociceptor neuron PVD, since a hypomorphic *zag-1* allele (*zd86*) led to the development of dendritic arbors in PVM (Smith, et al., 2013). Here, we found that a *zag-1* null mutation (*hd16*) led to developmental arrest at the L1 stage (AVM and PVM do not arise until late in the L1 stage) and the loss of expression of all the tested TRN fate markers in ALM and PLM neurons (Figure 3A-C). Thus, the role of ZAG-1 is important for general TRN fate specification.

We next tested the genetic interaction between *zag-1* and *mec-3* and found that the *zag-1 (hd16)* null allele did not affect the expression of the *mec-3* using a transcriptional reporter (Figure 3A), smFISH (Figure 3B), or MEC-3::GFP translational fusions (Figure 3C). In addition, expression of *zag-1(+)* from the *mec-3* promoter restored the expression of the TRN fate marker *mec-17p::RFP* in *zag-1* mutants, suggesting that ZAG-1 induces the expression of TRN terminal

differentiation genes cell-autonomously. Since *mec-3* expression is completely dependent on *unc-86*, we expected and found that the expression of *unc-86* was not changed in *zag-1* mutants.

Using a fosmid-based GFP translational fusion, we found that *zag-1* was expressed in the six TRNs but not FLP and PVD neurons (Figure 3D), and the expression of *zag-1* in TRNs was not affected by mutations in *mec-3* (Figure 3E). Therefore, *zag-1* and *mec-3* are transcriptionally independent of each other. We also failed to find any physical interaction between ZAG-1 and MEC-3 in yeast two-hybrid assays (Figure S1B-D). Together, our results suggest that ZAG-1 promotes TRN fate independently of the previously identified TRN fate determinants UNC-86 and MEC-3; the expression of the three transcription factors only overlap in the TRNs and thus can form a unique combinatorial code for TRN fate.

Smith *et al.* (2013) found that another conserved transcription factor AHR-1 (aryl hydrocarbon receptor) controls the differentiation of the anterior post-embryonic TRN neuron AVM; in *ahr-1* null mutants, AVM cells adopted a PVD-like multidendritic shape. We found that although expressed in all six TRNs, *ahr-1* is only required for the expression of TRN markers in AVM neurons but not in the other TRNs (ALMs, PLMs, and PVM), suggesting that AHR-1 plays a subtype-specific role in TRN fate specification (Figure S2A-C). Thus, our findings disagree with the hypothesis presented by Smith *et al.* (2013) that AHR-1 and ZAG-1 function in parallel to specify TRN morphology in the AVM and PVM, respectively. Instead, ZAG-1 is required for TRN fate adoption in general, whereas AHR-1 functions only in AVM. Moreover, Smith *et al*. (2013) found that AHR-1 specified TRN fate in AVM by regulating *mec-3*, whereas we found that ZAG-1 promoted TRN fate independently of *mec-3*.

ZAG-1 safeguards TRN fate by suppressing TRN fate inhibitors EGL-44/EGL-46

The TEA domain transcription factor EGL-44 and the Zn-finger protein EGL-46 inhibited the TRN fate in the FLP neurons (Wu, et al., 2001). These proteins appeared to work together to regulate gene expression and physically interacted in yeast two-hybrid assays (Figure S3; the EGL-44 and EGL-46 complex is hereafter abbreviated as EGL-44/EGL-46). Both genes were normally expressed in the FLP neurons, and the expression of *egl-46* was dependent on *egl-44* (Figure 4A and B). Our lab previously found and we confirmed that mutations in *egl-44* and *egl-46* caused the ectopic expression of *mec-17p::GFP* and other TRN reporters in FLP neurons (Figure 4A and C; Table S2). *egl-44* and *egl-46* were not expressed in differentiated TRNs (Figure 4A and B).

Since *zag-1* was selectively expressed in TRNs but not FLP neurons, we tested if ZAG-1 promoted the TRN fate by preventing the activation of EGL-44/EGL-46 in these cells. Consistent with this hypothesis, GFP reporters for both *egl-44* and *egl-46* were ectopically expressed in the TRNs in *zag-1* mutants (Figure 4D). In addition, *egl-44* mRNA, as measured by smFISH, was increased in *zag-1* TRNs, an indication that ZAG-1 transcriptionally repressed *egl-44* (Figure 4E). Importantly, *egl-44* and *egl-46* are epistatic to *zag-1*, since mutations in them restored the expression of TRN fate markers in *zag-1*-deficient animals (Figure 4C and Table S2). Thus, the TRN cell fate did not require ZAG-1 in the absence of EGL-44/EGL-46. Instead, ZAG-1 promoted the TRN fate through a double inhibition mechanism by preventing the expression of the TRN fate inhibitors *egl-44* and *egl-46*.

We also found that ZAG-1 loss affected general neurite growth and guidance in TRNs, since *egl-44; zag-1* and *zag-1; egl-46* mutants showed shortened and misguided TRN neurites. These results extend those of Smith *et al.* (2013), who found that ZAG-1 inhibited dendritic

growth in PVM neurons, and are consistent with the function of ZAG-1 in regulating axon guidance in other neurons (Clark and Chiu, 2003; Wacker, et al., 2003).

Misexpression of ZAG-1 converts FLP neurons into TRN-like cells

To address whether *zag-1* could affect FLP differentiation, we misexpressed *zag-1* in FLP neurons using the *mec-3* promoter. This misexpression led to the expression of *mec-17p::GFP* and other TRN markers (Figure 5A and Table S2) and diminished expression of *egl-44* and *egl-46* (Figure 5A and B) in FLP neurons. These data suggest that misexpressed ZAG-1 converts FLP neurons to a TRN-like fate by inhibiting *egl-44* and *egl-46*. The expression of *mec-3p::zag-1* transgene also activated TRN markers in PVD neurons (Figure 5A), which is consistent with previous findings that ZAG-1 prevents PVM neurons from adopting a PVD-like morphology (Smith, et al., 2013). Thus, ZAG-1 not only prevents PVM neurons from taking on the PVD fate but can also turn PVD neurons into TRN-like cells. Since *egl-44* was not expressed in PVD neurons, the misexpressed ZAG-1 presumably directs this conversion of cell fate by inhibiting some unidentified factor(s) that normally suppresses the TRN fate in PVD neurons.

We also expressed *ahr-1* from the *mec-3* promoter and found that misexpressed AHR-1 could activate TRN fate markers in PVD but not FLP neurons (Figure S2D-E). Overexpression of AHR-1 in TRNs also caused morphological defects, such as the growth of an ectopic ALM posterior neurite (Figure S2D), which does not occur when ZAG-1 was overexpressed. These results further support the hypothesis that ZAG-1 and AHR-1 have different functions.

EGL-44/EGL-46 prevents *zag-1* expression in FLP neurons

Given the mutually exclusive patterns of *zag-1* and *egl-44/egl-46* expression in FLP neurons and TRNs, we next tested whether EGL-44/EGL-46 regulated *zag-1* expression.

Mutations in *egl-44* and *egl-46* resulted in ectopic expression of a *zag-1::EGFP* reporter in FLP neurons. Misexpression of *egl-44* from the *mec-3* promoter is sufficient to suppress the TRN fate, because it activates the endogenous *egl-46*, a cofactor that is required for EGL-44 functions (Figure 5B and D; Wu, et al., 2001). We found that EGL-44 misexpression reduced, but did not completely eliminate, the expression of *zag-1* in TRNs (Figure 5B and C). Therefore, the positive TRN fate regulator ZAG-1 and the negative regulator EGL-44/EGL-46 reciprocally inhibit each other's expression, and their expression must be regulated so one or the other predominates.

Simultaneous misexpression of both EGL-44 and ZAG-1 from the *mec-3* promoter, however, led to the activation of TRN marker *mec-17p::GFP* in all *mec-3*-expressing neurons (TRNs, FLP, and PVD; Figure 5E), because ZAG-1-mediated repression of endogenous *egl-46* blocked the effects of EGL-44. In fact, supplying EGL-46 using the *mec-3* promoter removed the dominance of ZAG-1 and turned off the expression of TRN markers again (Figure 5F). Moreover, expression of an *egl-46*::GFP reporter was reduced by ZAG-1 despite the presence of misexpressed EGL-44 (Figure 5B), suggesting that ZAG-1 can repress *egl-46* transcription both through *egl-44* and independently of *egl-44*. Functionally, animals carrying the *uIs211[mec-3p::egl-44]* transgene were completely insensitive to gentle touch; co-expression of *zag-1* from *mec-3* promoter could restore the sensitivity in animals where the TRN markers were reactivated but failed to do so when *mec-3p::egl-46* was expressed (Figure S4A). Thus, the activity of EGL-44/EGL-46 complex predominates over ZAG-1 when both expressed from the same heterologous promoter in the same cell, leading to the suppression of TRN fate. The action of these competing inhibitors is not likely through direct interaction, since we did not observe any

physical interaction of ZAG-1 with either EGL-44 or EGL-46 or of MEC-3 with EGL-44, EGL-46, and ZAG-1 in yeast two-hybrid assays (Figure S3).

EGL-44/EGL-46 and ZAG-1 regulate a switch between FLP and TRN fates

We next investigated how EGL-44/EGL-46 and ZAG-1 regulated FLP fate using *sto-5* (Topalidou and Chalfie, 2011) and several other genes (*dma-1*, *bicd-1*, and *flp-4*) that were expressed in FLP neurons (Aguirre-Chen, et al., 2011; Liu and Shen, 2011; Kim and Li, 2004) but not the TRNs (Figure 6A). Expression of all four genes in the FLP neurons depended on EGL-44 and EGL-46 (Figure 6B and C). Moreover, the expression of the FLP and TRN markers were mutually exclusive; FLP neurons in *egl-44* and *egl-46* mutants or animals carrying *mec-3p::zag-1* transgene never showed mixed expression of the FLP- and TRN-specific genes (Figure 6E). Mutations in *zag-1* or the misexpression of *egl-44* in TRNs activates the FLP markers in addition to turning off the TRN markers (Figure 6B and S5), and we did not observe any mixed expression of both types of markers under those conditions (Figure 6E). Since FLP markers were not expressed in *egl-44; zag-1* double mutants, our results suggest that EGL-44/EGL-46 not only repressed the TRN fate but also promoted the FLP fate. Thus, the action of EGL-44/EGL-46 and ZAG-1 results in two mutually exclusive fates.

This bistable switch affected neuronal morphology as well as transcriptional reporters. For example, ALM neurons that assumed a FLP-like fate because of *egl-44* misexpression grew an ectopic posterior neurite (Figure S5A), and *zag-1*-expressing and thus TRN-like FLP and PVD neurons had far fewer dendritic branches than the wild-type cells (Figure S4B). This latter result is consistent with ZAG-1's role in preventing PVM from adopting a PVD-like morphology (Smith, et al., 2013).

dma-1, *bicd-1*, and *flp-4* were also expressed in PVD neurons, suggesting that FLP and PVD share a common genetic program. The expression of these PVD markers was blocked by the misexpression of ZAG-1 in PVD neurons, suggesting that ZAG-1 represses the unknown PVD fate selector. On the other hand, the *sto-5* reporter was expressed in FLP but not PVD neurons, indicating difference in the transcriptome of the two cell types (Figure S5E).

The activation of the FLP and PVD markers, like the TRN markers, required UNC-86 and MEC-3, which are the common terminal selectors for the three fates (Figure 6E and S5C). We identified one common terminal differentiation gene, *mec-19*, which was expressed in TRN, FLP, and PVD neurons at similar levels (and no other cells); its expression depended on UNC-86 and MEC-3 but was not affected by the loss of EGL-44/EGL-46 and ZAG-1 (Figure 6D-E). Thus, *unc-86* and *mec-3* serve as the "ground-state selectors" that promote a common ground states shared by the three types of neurons, and *mec-19* is a marker for this ground state. Subsequently, *zag-1* and *egl-44* act as "modulators" that shift the ground state towards specific fates. Therefore, the development of a particular fate requires the activities of both the ground-state selectors and the fate-restricting modulators.

EGL-44/EGL-46 inhibits the expression of TRN genes by competing with UNC-86/MEC-3 for DNA binding and by suppressing ALR-1

We next investigated the mechanism, by which EGL-44/EGL-46 prevented the expression of TRN terminal differentiation genes. Although we found two UNC-86/MEC-3 binding sites required for the expression of a minimal TRN-specific promoter (a 184-bp *mec-18* promoter) in TRNs, we failed to identify any additional *cis*-regulatory element that mediated its repression in FLP neurons (Figure S6A). Thus, EGL-44/EGL-46 seems unlikely to suppress TRN markers *via* distinct, repressive elements. We then tested the possibility that EGL-44/EGL-

46 acts through the UNC-86/MEC-3 binding site, since EGL-44 belongs to the TEA domain class transcription factors, which recognize DNA sequences similar to the UNC-86/MEC-3 binding site (Figure 7A) (Zhang, et al., 2002; Jiang, et al., 2000). EGL-44 and the EGL-44/EGL-46 complex bound to previously identified UNC-86/MEC-3 motifs in the *mec-4*, *mec-7*, *mec-17*, and *mec-18* promoters in electrophoretic mobility shift assays (Figure 7B and C). These results suggest that EGL-44 mediates the direct contact with the TRN promoters and EGL-46 acts as a corepressor. More importantly, EGL-44/EGL-46 outcompeted UNC-86/MEC-3 for the binding to these same sequences (Figure 7C), suggesting that the *cis*-regulatory sites normally bound by UNC-86/MEC-3 biochemically prefer EGL-44/EGL-46. Therefore, EGL-44/EGL-46 in the FLP neurons may prevent the activation of TRN genes by occluding the UNC-86/MEC-3 binding sites essential for the expression of TRN fate.

However, how EGL-44/EGL-46 avoids inhibiting UNC-86/MEC-3 targets that are commonly expressed in FLP and TRNs is unclear. One such target is *mec-3* itself, whose activation and maintenance depend on two UNC-86/MEC-3 binding sites (Xue, et al., 1992). We found that the two sites were not bound by EGL-44 (Figure 7B), which explains why *mec-3* expression is not affected by the presence of EGL-44/EGL-46 in FLP neurons. Comparing the *cis*-regulatory motif sequences on *mec-3* promoter with those on TRN-specific promoters, we found that the four nucleotides (positions 1 to 4 in Figure 7A) following the UNC-86/MEC-3 binding sequence showed significant divergence between EGL-44-binding and nonbinding sites. In particular, the EGL-44-binding sites all contain adenine at the fourth position and the nonbinding sites do not (Figure 7A). Changing this adenine to thymine eliminated the binding of EGL-44 to the site on *mec-4* promoter (Figure 7B). However, converting other nucleotides to adenine at the fourth position was not sufficient to enable the binding of EGL-44 to *mec-3*

promoter motifs and coordinated change of the nucleotides on the first and second positions are needed to evoke EGL-44 binding (Figure 7B). Similar results were also obtained for the UNC-86/MEC-3 target gene *mec-10*, which is expressed in both TRN and FLP cells (Huang and Chalfie, 1994). The UNC-86/MEC-3 site on *mec-10* promoter was not bound by EGL-44 and changing the nucleotides on the first and fourth position enabled EGL-44 binding. Thus, our results suggest that EGL-44/EGL-46 can differentiate TRN/FLP common genes from TRN-specific genes *via* the *cis*-regulatory sequences in their promoters.

We next tested whether converting the EGL-44 binding site to a nonbinding site in a TRN promoter could prevent the suppression by EGL-44/EGL-46 in FLP neurons *in vivo*. Contrary to our expectation, a *mec-4* promoter reporter harboring the adenine-to-thymine change at the fourth position was not activated in FLP neurons, although its TRN expression was preserved. In the 184-bp *mec-18* promoter, changing the four nucleotides following the UNC-86/MEC-3 binding sequence from ACCA to CTAT completely abolished EGL-44 binding *in vitro*; but the mutant reporter was only weakly expressed in ~30% of FLP neurons and remained silenced in the rest ~70% (Figure S6B-D). In addition, creating an EGL-44-binding site in the *mec-3* promoter by mutating the regulatory motif did not suppress *mec-3* expression in FLP neurons (Figure S6B). The discrepancy between the *in vitro* and *in vivo* results suggests that the lack of EGL-44 binding to TRN differentiation genes *per se* is not sufficient to distinguish TRN fate from the TRN/FLP ground state. In addition to directly binding to the *cis*-regulatory elements of TRN promoters, EGL-44/EGL-46 may also inhibit TRN genes by activating or repressing the expression of other *trans*-acting factors.

One such factor may be ALR-1, which is an ortholog of human Arx and Drosophila *aristaless* and is needed for the robust differentiation of TRN fate (Topalidou, et al., 2011). Loss

of *alr-1* variably reduced but did not eliminate the expression of TRN markers in the TRNs in wild type and strongly eliminated the ectopic expression of TRN markers in FLP neurons in *egl-44* and *egl-46* mutants (Figure 7D and S6E). This difference between reduction and elimination may result from the fact that FLP neurons have lower *mec-3* expression than the TRNs (Topalidou, et al., 2011) and thus a stronger need for ALR-1 to activate TRN markers. The observation that *alr-1* is epistatic to *egl-44* and *egl-46* suggested that ALR-1 is a downstream effector suppressed by EGL-44/EGL-46. Indeed, a fosmid-based *alr-1* translational reporter, which was normally expressed in TRNs but not FLP neurons, became de-repressed in FLP cells in *egl-44* and *egl-46* mutants (Figure 7D and S6E).

The lack of ALR-1 and the lower level of *mec-3* in wild-type FLP neurons may explain the inactivation of the mutated TRN fate reporters that cannot be bound by EGL-44/EGL-46. Supporting this hypothesis, forced expression of ALR-1 in FLP neurons ectopically activated even the wild-type TRN fate reporters (Figure 7D) (Topalidou, et al., 2011), suggesting that ALR-1 not only promotes TRN fate but can also overcome the direct suppression from EGL-44/EGL-46 on the TRN genes. This ability of ALR-1 may result from ALR-1 upregulating *mec-3* (Topalidou, et al., 2011) and directly interacting with TRN promoters (ChIP-seq data; Figure S6F). Therefore, EGL-44/EGL-46 inhibits TRN-specific genes by both occupying the UNC-86/MEC-3 site in the TRN promoters and suppressing TRN fate-promoting transcription factor ALR-1. Consistent with this model, overexpression of MEC-3 in FLP neurons could overcome EGL-44/EGL-46-mediated inhibition and activate the TRN program (Topalidou and Chalfie, 2011) presumably by both retaking the UNC-86/MEC-3 sites and by activating *alr-1*, which is a *mec-3*-dependent gene (Topalidou, et al., 2011).

Spatial and temporal expression of *zag-1* and *egl-44* are mutually exclusive

Given the mutual inhibition of *zag-1* and *egl-44* in the FLP neurons and the TRNs, we asked whether the two genes were generally expressed in different cells in the nervous system. Using the neurotransmitter maps for glutamatergic, cholinergic, and GABAergic neurons (Gendrel, et al., 2016; Pereira, et al., 2015; Serrano-Saiz, et al., 2013), we found that, in addition to TRNs, *zag-1* was expressed in the AIB, AIM, AIN, AIZ, AVA, AVB, AVD, AVE, AVG, AVK, AVL, M4, M5, RIA, RIB, RIF, RIG, RIM, RIV, RMD, RME, RMF, RMH, SIA, and SMD neurons in the head, all the DD, VD, and VC neurons in the ventral cord, and the DVA, DVB, LUA, PDA, PVC, PVP, PVQ, PVR, and PVT neurons in the tail. *zag-1* is also expressed in the serotonergic HSN neurons. In comparison, *egl-44* expression was much more restricted, being found only in the FLP, ADL, and SAB neurons, as well as a few VA and VB motor neurons. Moreover, *egl-44* was widely expressed in hypodermis, pharynx, and intestine, whereas *zag-1* expression was absent in these tissues. We also constructed an *egl-44::RFP* reporter and cross it with *zag-1:EGFP* and did not observe overlapping expression in any cell (Figure S7A). The above data support the hypothesis that *zag-1* and *egl-44* expression are mutually exclusive in the nervous system and throughout all tissues of the animal. Given the mutual inhibition of EGL-44 and ZAG-1 in the TRNs/FLPs, we expected that the loss of *zag-1* might affect *egl*-44 expression more broadly. We did not, however, observe a systematic upregulation of EGL-44::GFP expression in the nervous system in either the *zag-1(zd86)* hypomorphic mutants or the arrested L1 animals of *zag-1(hd16)* null mutants (Figure S7B and C). Similarly, tissues like hypodermis, pharynx and intestine did not gain *zag-1* expression upon the loss of *egl-44* either, suggesting that, in addition to the reciprocal inhibition, other activating signals are needed to create the expression pattern of the two transcription factors.

A few neurons expressed both *zag-1* and *egl-44* but did so at different times. *egl-44* and *egl-46* were expressed in the precursors of the postembryonic TRNs (the AVM and PVM neurons) and persisted in these cells for a few hours after their generation to promote cell cycle exit (Feng, et al., 2013; Wu, et al., 2001), but they were not expressed in terminally differentiated AVM and PVM. These differentiated cells expressed *zag-1*, which promoted the TRN fate. Similarly, *egl-44* and *egl-46* were transiently expressed in the early embryos in HSN neurons before they migrated and differentiated (Wu, et al., 2001) (Figure S7D), whereas *zag-1* was expressed in the terminally differentiated HSN neurons in adults and was required for the activation of the HSN fate marker *tph-1p::GFP* (Figure S7E and F).

Discussion

RNAi screen is a systematic method to identify genes involved in cell fate determination

We demonstrate here that a systematic RNAi screen of a library of transcription factors can be used to identify neuronal cell fate regulators, particularly genes whose mutations lead to lethality or sterility. Our previous forward genetic screens, which searched for viable mutants with touch-sensing defects, despite reaching saturation, only identified *unc-86* and *mec-3* as the TRN fate determinants (Chalfie and Au, 1989; Chalfie and Sulston, 1981). Using the RNAi screen, we not only recovered *unc-86* and *mec-3* blindly, but also identified *ceh-20*, *ldb-1*, and *zag-1* as genes required for the expression of TRN fate. These latter genes would not have been identified in our previous screens, because null mutations in them lead to early larval arrest. Moreover, our previous neuronal RNAi screen of nonviable genes would not have identified these genes either, because it only looked for postembryonic effects on touch sensitivity after TRN differentiation was completed (Chen, et al., 2015). In contrast, the current RNAi screen

knocked down gene expression in the embryos and could identify genes whose silencing perturbed TRN fate specification.

Although previous studies of individual genes identified *ceh-20* (Zheng, et al., 2015) and *ldb-1* (Cassata, et al., 2000) as regulators of TRN fate, RNAi screens, such as we have done, provide a systematic, complementary approach to identify genes controlling neuronal differentiation. *ceh-13* was also known to affect the TRN fate in ALM neurons (Zheng, et al., 2015) but was not identified from the RNAi screen, suggesting the existence of false negatives. Importantly, through the screen, we identified ZAG-1 as a new TRN fate determinant and the third piece in a combinatorial code (UNC-86, MEC-3, and ZAG-1) that defines TRN fate. Although previous studies suggested that ZAG-1 regulates PVM morphological differentiation using a viable hypomorphic *zag-1* allele (Smith, et al., 2013), our RNAi screen and the subsequent studies using a null allele found that ZAG-1 is essential for the acquisition of TRN fate more generally.

Our RNAi results, however, raise two concerns about the reliability and applicability of such screens. First, only 5 of the 14 genes we identified from the screen were confirmed with mutants. The remaining 9 genes (65%) appeared to be false positives, since their null mutations did not result in TRN differentiation defects, although RNAi treatment against these genes all repeatedly showed the loss of TRN marker expression. These false positives probably resulted from mistargeting of the dsRNA. A similar problem was encountered in a genome-wide RNAi screen for genes involved in the specification of ASE neuron in *C. elegans* (Poole, et al., 2011), since a significant fraction of the genes identified from the screen could not be verified using mutants (Oliver Hobert, personal communication). Second, although both ALM and PLM neurons required similar factors for cell fate determination, RNAi against *unc-86* and the other

fate regulators caused much stronger effects in the ALM neurons than in PLMs. This bias towards the anterior was previously observed in our enhanced neuronal RNAi system (Calixto, et al., 2010), suggesting that neurons with the cell body located in the tail region, especially positioned posteriorly to the anus, may be less affected by dsRNA introduced by feeding than those located in the anterior half of the animal. This bias led to the failure of recovering *egl-5*, whose loss was known to only affected PLM differentiation (Zheng, et al., 2015).

Non-selector type of cell fate regulators promotes fate acquisition indirectly

The key concept of "terminal selectors" is that transcription factors controlling neuronal cell fate directly activate terminal differentiation genes associated with the identity and function of the neuron through common *cis*-regulatory elements (Hobert, 2011; Hobert, et al., 2010). POU homeodomain protein UNC-86 and LIM homeodomain protein MEC-3 are examples of such terminal selectors, since they directly activate TRN genes *via* the UNC-86/MEC-3 binding sites in the promoter of TRN genes (Duggan, et al., 1998).

Here, however, we found that at least two differentiation regulators that are required for TRN fate indirectly promote the expression of TRN genes. The LIM domain-binding protein LDB-1 binds to MEC-3, which protects MEC-3 from degradation and enables it to activate downstream TRN genes. ZAG-1 promotes the expression of TRN markers by repressing the cell fate inhibitors EGL-44 and EGL-46. Although both LDB-1 and ZAG-1 are absolutely required for the adoption of TRN fate, neither of them directly associated with the DNA elements of TRN genes and control the activity of TRN promoters. ZAG-1 belongs to the ZEB family of transcription factors, which mostly function as transcriptional repressors (Vandewalle, et al., 2009); so, it could only indirectly induce the expression of TRN genes. Moreover, unlike the highly confined expression of classical terminal selectors (e.g. *mec-3* is only expressed in ten

neurons), LDB-1 and ZAG-1 were expressed broadly. *ldb-1* is widely expressed in the nervous system, the reproductive system, and the muscles (Cassata, et al., 2000), and *zag-1* is expressed in many neurons in the head and tail ganglia and ventral cord motor neurons, as well as various muscles (Wacker, et al., 2003). These features suggest that LDB-1 and ZAG-1 may represent a non-selector type of cell fate regulators that is phenotypically identical to but mechanistically different from the conventional terminal selectors.

A bistable negative feedback loop controls binary cell fate decisions

A central question for neuronal differentiation is how the extraordinary diversity of neuron types are generated and maintained. The "combinatorial code" hypothesis suggests that the identity of a neuron type or subtype is determined by a combination of transcription factors (Allan, et al., 2003; Jessell, 2000; Mitani, et al., 1993), but how their actions are coordinated mechanistically in a regulatory network is not well understood. One way to delineate the activities of a complex code is to investigate how binary fate decisions are made to distinguish one neuron type from another. In this study, we found that within the three-factor code for the TRN fate, two factors (UNC-86 and MEC-3) form a selector for a ground state shared by FLP and TRNs, whereas the third factor (ZAG-1) controls a binary choice between the two fates.

Although FLP and TRN share the same selectors, FLP neurons fail to express the TRN genetic program because EGL-44/EGL-46 blocks the activation of TRN genes by occupying the *cis*-regulatory elements bound by UNC-86/MEC-3 and by repressing the TRN fate regulator ALR-1 (Figure 8). This inhibition of TRN fate is very robust, since EGL-44/EGL-46 outcompete UNC-86/MEC-3 for the binding of the same *cis*-regulatory element in TRN promoters (Figure 7C) and misexpressing *egl-44* in TRNs can completely shut off the expression of TRN genes (Figure 5D). As a consequence, TRNs utilize a protective mechanism that uses ZAG-1 to

strongly repress both *egl-44* and *egl-46* and thus ensures their absence in the TRNs. Reciprocally, EGL-44/EGL-46 also silences *zag-1* in FLP neurons, which secured the expression of *egl-44* and *egl-46* and prevented the adoption of TRN fate. Such mutual repression enables a bistable switch that locks the cell fate in one of the two possible states (Figure 8). In the absence of such a switch (e.g. in *egl-44; zag-1* double mutants), both TRNs and FLP neurons expressed the TRN genetic program (Figure 4), suggesting that the ground state is a TRN-like fate and the selectors UNC-86/MEC-3 activate TRN genes by default. EGL-44/EGL-46-induced modification of the default genetic program gives rise to the FLP fate.

Interestingly, the Drosophila homolog of ZAG-1, Zfh1, also regulates binary cell fate choices. Zfh1 promotes a GW motor neuron fate over an EW interneuron fate in the 7-3 neuroblast lineage, and its expression is suppressed by the steroid receptor family transcription factor Eagle in the EW interneuron (Lee and Lundell, 2007); whether Zfh1 also suppressed Eagle in GW motor neuron is unclear. In another example, the mutual antagonism between Zfh1 and another zinc-finger transcription factor, Lame duck (Lmd), regulates the decision between pericardial cell and fusion competent myoblast fates in the mesoderm (Sellin, et al., 2009). Although a genetic interaction between Zfh1 and the Drosophila EGL-44, Scalloped, has not been reported, these results suggest that Zfh1 can form regulatory switches with other transcription factors. Mouse homologs of ZAG-1, ZEB1 and ZEB2, repress tissue differentiation during early embryogenesis, induce epithelial mesenchymal transition (EMT), and are essential for neural tube development (Vandewalle, et al., 2009). Their roles in the specifying terminal cell fates, however, are unclear, because their knockout leads to embryonic lethality (Miyoshi, et al., 2006).

The choosing of one of two alternative fates appears to be a common way to diversify neuronal types and subtypes during development, as shown by several examples in *C. elegans* and Drosophila (Sarin, et al., 2007; Mikeladze-Dvali, et al., 2005). In vertebrates, neuronal fate, especially neurotransmitter identity, is also subjected to such binary regulation, but the underlying mechanism is largely unknown. For example, *Lbx1* and *Helt* promotes GABAergic identity and suppresses glutamatergic differentiation in the dorsal spinal cord (Cheng, et al., 2005) and the mesencephalon (Nakatani, et al., 2007), respectively. The selector gene *Fezf2* induces the opposite choice in mouse corticospinal motor neurons by activating *Vglut1* to produce glutamatergic cells and by repressing *Gad1* to inhibit a GABAergic fate (Lodato, et al., 2014). However, whether the two fates are exclusive to each other because of a bistable switch is unknown. One important feature of those binary fate choices is that the selector often simultaneously induces one fate and suppresses the other. Whether there two activities can be uncoupled is unclear. Our studies of the FLP versus TRN fate choice suggests that uncoupling is possible. Although EGL-44/EGL-46 simultaneously induces FLP genes and suppresses TRN genes, ZAG-1 appears to be solely devoted to inhibiting FLP fate (and possibly PVD fate) and is not involved in directly activating TRN genes. The uncoupling between UNC-86/MEC-3 activating TRN fate and ZAG-1 preventing alternative fates suggest that the two modules may be independently evolved.

Binary fate switches may enable the generation of neuronal diversity

Because the introduction of self-reinforcing, bistable switches can generate diversity within pre-existing neuronal fates, we envision that the extraordinary variety of neuronal types in the nervous system evolved from a few primitive neuronal fates through stepwise addition of binary switches. For example, since TRN, FLP, and PVD fates all require the same selectors,

UNC-86 and MEC-3, they may be derived from a common ancestral fate. In fact, the existence of UNC-86/MEC-3 target genes like *mec-19*, which is expressed in all the FLP, PVD, and TRNs but not any other neuron, suggests such an ancestral fate.

One possible evolutionary derivation of these cells is that some of the ancestral TRN-like cells acquired the ability to express *egl-44* and *egl-46*. This expression led to the suppression of TRN genes, the activation of FLP genes and the emergence of this new cell type. The fact that the ground state is a TRN-like state seems to support this hypothesis. Moreover, since *egl-44* is mostly expressed in non-neuronal tissues, mutations in the regulatory elements of *egl-44* might have allowed expression in some neurons. Thus, EGL-44 may have been co-opted to induce divergence among neurons that share a common fate, and this cell fate divergence is subsequently stabilized by the establishment of a negative feedback loop between EGL-44 and ZAG-1.

Alternatively, a regulatory element in the *zag-1* gene, which was already present in the ancestral TRNs to repress *egl-44* and *egl-46*, may have been mutated to prevent its expression in some ancestral TRN cells. This loss led to the de-repression of *egl-44* and *egl-46* and the subsequent acquisition of the FLP fate. The widespread expression of *zag-1* in neurons suggests that it may be in general required for the differentiation of many neuronal fates, and selective loss of *zag-1* expression in some cells through changes in *cis*-regulatory sequences may prevent these cells from committing to particular fates and allowing them to adopt alternative ones.

Overall, we imagine that a limited number of ground-state selectors may first define a handful of shared, and subsequently a series of binary fate switches carry out further differentiation that modifies the ground state to generate a diverse array of terminal neuronal fates. The broad expression of ZAG-1 in the nervous system and the lack of increased EGL-44

expression in *zag-1* mutants suggests that ZAG-1 may form bistable regulatory loops with many different inhibitors. In fact, ZAG-1 is required for the fate specification of at least TRN, HSN, and PVQ neurons (this study; Clark and Chiu, 2003), although whether ZAG-1 also regulates HSN and PVQ fates by repressing alternative fates is unclear. Moreover, the fact that ZAG-1 expression is only found in a subset of neurons that express the same selector supports the hypothesis that ZAG-1 may ensure diversification from a shared ground state through down-regulation of inhibitors; e.g. *zag-1* is expressed in 7/17, 6/15, and 3/11 classes of neurons that use UNC-86, UNC-3, and CEH-14 as a terminal selector, respectively (classification according to Hobert, 2016). Finally, the broad expression of ZEB family transcription factors in the nervous system of both Drosophila (Lai, et al., 1991) and mice (Vandewalle, et al., 2009) suggests possibly a conserved role for them in regulating binary fate choices and the generation of neuronal diversity.

Author Contributions

C.Z., F.Q.J., and B.L.T. performed experiments and analyzed the data. C.Z. and M.C. conceived the study and wrote the manuscript. M.C. supervised the work and provided funding. F.Q.J. and B.L.T. contributed equally to the work.

Acknowledgement

We thank Songtao Jia and Elizabeth Miller for sharing materials and reagents. We also thank Oliver Hobert, Richard Mann, and the members of our laboratory for helpful discussions and comments. This work was supported by NIH grants GM30997 and GM122522 to MC.

STAR Methods

❖ Strains, Constructs, and transgenes

C. elegans strains and mutant alleles

C. elegans wild type (N2) and mutant strains were maintained as previously described (Brenner, 1974). Most strains were provided by the *Caenorhabditis* Genetics Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440), or the National BioResource Project of Japan. VH514, *zag-1 (hd16) /unc-17 (e113) dpy-13 (e184)* and VC639, *ldb-1 (ok896)/szT1; +/szT1[lon-2 (e678)]* were used as the balanced null alleles for *zag-1* and *ldb-1*, respectively. *zag-1 (zd86)* was used as a hypomorphic allele for *zag-1*. For other genes identified from the RNAi screen (see below), we tested *zip-4 (tm1359)*, *nhr-119 (gk136908)*, *nhr-166 (gk613)*, *nhr-159 (tm2323)*, *egl-38 (ok3510)/nT1[qIs51]*, *hmbx-1 (ok3467)*, *fkh-2 (ok683)*, *lin-40 (ku285)*, *lin-40 (s1506) unc-46 (e177)/eT1*, *elt-6 (gk723)*, and *elt-6 (gk754)*. Other mutant alleles used in this study include *ahr-1(ju145)*, *egl-44(n1080)*, *egl-46(n1127)*, *mec-3(u184)*, and *unc-86(u5)*.

Constructs

A 2.4 kb *mec-3* promoter, a 5.9 kb *ldb-1* promoter, a 2.2 kb *zag-1* promoter, a 2.2 kb *bicd-1* promoter, 4.9 kb *dma-1* promoter, 3.1 kb *flp-4* promoter, 2.3 kb *sto-5* promoter, and 1.3 kb *mec-19* promoter were cloned from wild type (N2) genomic DNA into the Gateway pDONR221 P4-P1r vector. The genomic coding region of GFP, *ldb-1a*, *zag-1*, *egl-44*, *egl-46*, *alr-1*, and *ahr-1* were cloned into Gateway pDONR221. The resulted entry vectors, together with pENTR-*unc-54*-3’ UTR and the destination vector pDEST-R4-R3 were used in the LR reaction

to create the final expression vectors. Gateway cloning was performed according to the manual provided by Life Technologies (Grand Island, NY).

DNA constructs TU#625 and TU#626 (Wu, et al., 2001) contain translational GFP fusion of *egl-44* and *egl-46*, respectively, and were injected into animals to form reporters for the two genes, named as *uIs215[egl-44::GFP]* and *uEx927[egl-46::GFP]*. TU#924 contains a 400 bp *mec-18* promoter inserted into pPD95.75 between HindIII and BamHI sites, and this *mec-18p::GFP* construct was used as a template to create a series promoter variants shown in Figure S10 using the Q5 site-directed mutagenesis kit from New England Biolabs (Ipswich, MA).

Transgenes

Transgenes *zdIs5[mec-4p::GFP]* I, *muIs32[mec-7p::GFP]* II, *uIs31[mec-17p::GFP]* III, *uIs115[mec-17p::RFP]* IV, *uIs134[mec-17p::RFP]* V, and *uIs72[mec-18p::mec-18::GFP]* were used as fluorescent markers for the TRN cell fate. *uEx1104[Pbicd-1::GFP]*, *uEx1105[dma-1p::GFP]*, *uEx1106[flp-4p::GFP]*, and *uIs232[sto-5p::GFP]* served as FLP fate markers. *zdIs13[tph-1p::GFP]* and *vsIs97[tph-1p::DsRed2]* were used as HSN fate marker. *uIs22[mec-3p::GFP]* and *uIs152[mec-3p::RFP]* were used as *mec-3* transcriptional reporter; *uEx1007[mec-3p::mec-3::GFP]* and *wgIs55[mec-3::TY1::EGFP::3xFLAG]* were used as *mec-3* translational reporter. *wgIs83[zag-1::TY1::EGFP::3xFLAG]*, *wgIs476[unc-86::TY1::EGFP::3xFLAG]*, *uEx1006[ldb-1p::GFP]*, *wgIs200[alr-1::TY1::EGFP::3xFLAG]*, *leEx1709[ahr-1::GFP]* and *uEx1107[mec-19p::GFP]* served as the reporters for *zag-1*, *unc-86*, *ldb-1*, *alr-1*, *ahr-1*, and *mec-19* respectively. *uIs211[mec-3p::egl-44]*, *uEx926[mec-3p::zag-1]*, and *uEx1027[mec-3p::alr-1]* were used for misexpression.

To perform cell identification, we crossed *wgIs83[zag-1::TY1::EGFP::3xFLAG]* and *uIs215[egl-44::GFP]* into the *otIs518[eat-4::SL2::mCherry::H2B]*, *otIs544[cho-*

1::SL2::mCherry::H2B], and *otIs564 [unc-47::SL2::H2B::mChopti]*, labeling glutamatergic, cholinergic, and GABAergic neurons (Gendrel, et al., 2016; Pereira, et al., 2015; Serrano-Saiz, et al., 2013), respectively. We identified *zag-1* and *egl-44* expressing neurons, based on the position and neurotransmitter identity of the cells expressing GFP.

❖ **RNAi screen**

RNAi screen was performed using a modified bacteria-feeding protocol previously reported (Poole, et al., 2011; Kamath, et al., 2003). We used the Ahringer RNAi library from Source Bioscience (http://www.lifesciences.sourcebioscience.com/) and the list of 392 RNAi clones targeting transcription factors were generated by searching WormBase WS238 using Gene Ontology terms related to "DNA binding" and "transcription factor activity" (see the complete list in Table S1). To perform the RNAi experiments, we seeded bacteria expressing dsRNA on NGM agar plates containing 6 mM IPTG and 100 µg/ml ampicillin. One day later, eggs from TU4429, *eri-1 (mg366); lin-15B (n744); uIs134[mec-17p::RFP]* animals were placed onto these plates; the eggs hatched and grew to adults at 20°C. The F1 progeny of these worms were scored for the expression of TRN markers at the second larval stage. RNAi clones were considered to be positive if more than 15% ($n > 20$) of the treated animals failed to show RFP expression in the ALM neurons in at least two of the three replicate plates. Three initial rounds of screens were performed on all the 392 clones, and 14 clones were found positive in all the three rounds. Two more screening rounds were then conducted on these 14 RNAi clones, which were all confirmed to be positive. We sequenced the inserts of all positive clones to confirm the identity of the target genes.

❖ **Yeast two-hybrid assay**

Yeast media and plates were prepared according to recipes from Clontech (Mountain View, CA) and yeasts were grown at 30°C. The yeast strain PJ69-4a (provided by Songtao Jia at Columbia University) used for the two-hybrid assays contains GAL1-HIS3, GAL2-ADE2, and GAL7-lacZ reporters. Vectors pGAD424 and pGBT9 (Clontech) were used to express proteins fused to the yeast activating domain (AD) and binding domain (BD), respectively. cDNA fragments of *mec-3*, *unc-86*, *zag-1*, *ldb-1*, *egl-44*, and *egl-46* were cloned into the two-hybrid vectors either using restriction enzymes or with Gibson Assembly (NEB).

Combinations of the AD or BD vectors were co-transformed into yeast using the Frozen-EZ II kit from Zymo Research (Irvine, CA) and using empty vectors as negative controls. Growth assays were performed by growing individual colonies overnight in selective media lacking tryptophan and leucine. Cultures were then diluted to let OD600 become 0.5, and 10 μl of a further 1:10 diluted culture were spotted onto plates lacking histidine to test the expression of the HIS3 reporter. Plates were imaged after 2 days of growth. Liquid β-galactosidase assays were performed using the Yeast β-Galactosidase Assay Kit (Thermo Scientific, Rockford, IL).

Electrophoretic mobility shift assay (EMSA)

Recombinant GST::EGL-44 proteins were produced in *E. coli* BL21 (DE3) using the expression vector pGEX-6p-1 (Amersham Pharmacia Biotech, UK) and purified using affinity chromatography columns filled with Glutathione Sepharose 4B beads (Amersham). EGL-44 was cleaved off the column using PreScission Protease (Amersham). EGL-46 was expressed using the pET32a vector (Novagen, Madison, WI) and purified using the S-Tag rEK purification kit (Novagen). UNC-86 and MEC-3 were produced according to previous methods(Xue, et al., 1993).

Gel mobility shift assays were performed using a modified protocol previously reported (Xue, et al., 1993). DNA probes were labeled with Biotin by Biotin 3' End labeling kit (Pierce, Rockford, IL) and then annealed into double strands. 100 ng of proteins were incubated with 0.005 ng probe at room temperature for 30 min, and the mixture was loaded onto a 10% TBE polyacrylamide mini gel (Bio-Rad, Hercules, CA). The gel was transferred to a 0.2 mm nylon membrane, which was then treated with UV light to crosslink DNA and proteins. Biotin-labeled DNA was detected using LightShift chemiluminescence EMSA kit (Pierce). The sequences of the probes are listed in Table S3.

❖ **smFISH, phenotypic scoring, and statistical analysis**

smFISH

Single-molecule fluorescence in situ hybridization (smFISH) was performed as described previously (Topalidou, et al., 2011). Imaging was conducted on a Zeiss Axio Observer Z1 inverted microscope with a CoolSNAP HQ2-FW camera (Photometrics, Tucson, AZ).

Phenotypic scoring

To examine the expression pattern of TRN markers, we grew animals at 20° C, examined them using the same microscope and recorded the percentages of TRN cells that express the fluorescent reporter in three independent experiments. The results are presented as aggregates; no significant differences were seen between replicates. To test transgenic animals, we injected DNA constructs (5 ng/μl for each expression vector) into the animals to establish stable lines carrying the extrachromosomal array; at least three independent lines were tested. In some cases, the transgene was integrated into the genome using γ-irradiation (Mello, et al., 1991), and at least three integrant lines were outcrossed and examined.

Statistical analysis

Statistical significance was determined using the Student's t-test for the majority of comparisons of two sets of data. For multiple comparisons, the Holm-Bonferroni method was used to correct the *p* values.

Reference

Aguirre-Chen, C., Bulow, H.E., and Kaprielian, Z. (2011). C. elegans bicd-1, homolog of the Drosophila dynein accessory factor Bicaudal D, regulates the branching of PVD sensory neuron dendrites. Development 138, 507-18.

Allan, D.W., St Pierre, S.E., Miguel-Aliaga, I., and Thor, S. (2003). Specification of neuropeptide cell identity by the integration of retrograde BMP signaling and a combinatorial transcription factor code. Cell 113, 73-86.

Bach, I., Carriere, C., Ostendorff, H.P., Andersen, B., and Rosenfeld, M.G. (1997). A family of LIM domain-associated cofactors confer transcriptional synergism between LIM and Otx homeodomain proteins. Genes Dev 11, 1370-80.

Blackshaw, S., Fraioli, R.E., Furukawa, T., and Cepko, C.L. (2001). Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. Cell 107, 579-89.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics 77, 71-94.

Calixto, A., Chelur, D., Topalidou, I., Chen, X., and Chalfie, M. (2010). Enhanced neuronal RNAi in C. elegans using SID-1. Nat Methods 7, 554-9.

Cassata, G., Rohrig, S., Kuhn, F., Hauri, H.P., Baumeister, R., and Burglin, T.R. (2000). The Caenorhabditis elegans Ldb/NLI/Clim orthologue ldb-1 is required for neuronal function. Dev Biol 226, 45-56.

Chalfie, M., and Au, M. (1989). Genetic control of differentiation of the Caenorhabditis elegans touch receptor neurons. Science 243, 1027-33.

Chalfie, M., and Sulston, J. (1981). Developmental genetics of the mechanosensory neurons of Caenorhabditis elegans. Dev Biol 82, 358-70.

Chatzigeorgiou, M., and Schafer, W.R. (2011). Lateral facilitation between primary mechanosensory neurons controls nose touch perception in C. elegans. Neuron 70, 299-309.

Chatzigeorgiou, M., Yoo, S., Watson, J.D., Lee, W.H., Spencer, W.C., Kindt, K.S., Hwang, S.W., Miller, D.M., 3rd, Treinin, M., Driscoll, M., et al. (2010). Specific roles for DEG/ENaC and TRP channels in touch and thermosensation in C. elegans nociceptors. Nat Neurosci 13, 861-8.

Chen, X., Cuadros, M.D., and Chalfie, M. (2015). Identification of nonviable genes affecting touch sensitivity in Caenorhabditis elegans using neuronally enhanced feeding RNA interference. G3 (Bethesda) 5, 467-75.

Cheng, L., Samad, O.A., Xu, Y., Mizuguchi, R., Luo, P., Shirasawa, S., Goulding, M., and Ma, Q. (2005). Lbx1 and Tlx3 are opposing switches in determining GABAergic versus glutamatergic transmitter phenotypes. Nat Neurosci 8, 1510-5.

Clark, S.G., and Chiu, C. (2003). C. elegans ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. Development 130, 3781-94.

Corbo, J.C., Lawrence, K.A., Karlstetter, M., Myers, C.A., Abdelaziz, M., Dirkes, W., Weigelt, K., Seifert, M., Benes, V., Fritsche, L.G., et al. (2010). CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. Genome Res 20, 1512-25.

Duggan, A., Ma, C., and Chalfie, M. (1998). Regulation of touch receptor differentiation by the Caenorhabditis elegans mec-3 and unc-86 genes. Development 125, 4107-19.

Etchberger, J.F., Lorch, A., Sleumer, M.C., Zapf, R., Jones, S.J., Marra, M.A., Holt, R.A., Moerman, D.G., and Hobert, O. (2007). The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron. Genes Dev 21, 1653-74.

Feng, G., Yi, P., Yang, Y., Chai, Y., Tian, D., Zhu, Z., Liu, J., Zhou, F., Cheng, Z., Wang, X., et al. (2013). Developmental stage-dependent transcriptional regulatory pathways control neuroblast lineage progression. Development 140, 3838-47.

Finney, M., and Ruvkun, G. (1990). The unc-86 gene product couples cell lineage and cell identity in C. elegans. Cell 63, 895-905.

Freyd, G., Kim, S.K., and Horvitz, H.R. (1990). Novel cysteine-rich motif and homeodomain in the product of the Caenorhabditis elegans cell lineage gene lin-11. Nature 344, 876-9.

Gendrel, M., Atlas, E.G., and Hobert, O. (2016). A cellular and regulatory map of the GABAergic nervous system of C. elegans. Elife 5.

Gupta, B.P., Wang, M., and Sternberg, P.W. (2003). The C. elegans LIM homeobox gene lin-11 specifies multiple cell fates during vulval development. Development 130, 2589-601.

Hobert, O. (2011). Regulation of terminal differentiation programs in the nervous system. Annu Rev Cell Dev Biol 27, 681-96.

Hobert, O. (2016). A map of terminal regulators of neuronal identity in Caenorhabditis elegans. Wiley Interdiscip Rev Dev Biol 5, 474-98.

Hobert, O., Carrera, I., and Stefanakis, N. (2010). The molecular and gene regulatory signature of a neuron. Trends Neurosci 33, 435-45.

Hsiau, T.H., Diaconu, C., Myers, C.A., Lee, J., Cepko, C.L., and Corbo, J.C. (2007). The cis-regulatory logic of the mammalian photoreceptor transcriptional network. PLoS One 2, e643.

Huang, M., and Chalfie, M. (1994). Gene interactions affecting mechanosensory transduction in Caenorhabditis elegans. Nature 367, 467-70.

Jessell, T.M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat Rev Genet 1, 20-9.

Jiang, S.W., Trujillo, M.A., Sakagashira, M., Wilke, R.A., and Eberhardt, N.L. (2000). Novel human TEF-1 isoforms exhibit altered DNA binding and functional properties. Biochemistry 39, 3505-13.

Jurata, L.W., Kenny, D.A., and Gill, G.N. (1996). Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development. Proc Natl Acad Sci U S A 93, 11693-8.

Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., et al. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature 421, 231-7.

Kaplan, J.M., and Horvitz, H.R. (1993). A dual mechanosensory and chemosensory neuron in Caenorhabditis elegans. Proc Natl Acad Sci U S A 90, 2227-31.

Kim, K., and Li, C. (2004). Expression and regulation of an FMRFamide-related neuropeptide gene family in Caenorhabditis elegans. J Comp Neurol 475, 540-50.

Lai, Z.C., Fortini, M.E., and Rubin, G.M. (1991). The embryonic expression patterns of zfh-1 and zfh-2, two Drosophila genes encoding novel zinc-finger homeodomain proteins. Mech Dev 34, 123-34.

Lee, H.K., and Lundell, M.J. (2007). Differentiation of the Drosophila serotonergic lineage depends on the regulation of Zfh-1 by Notch and Eagle. Mol Cell Neurosci 36, 47-58.

Liu, O.W., and Shen, K. (2011). The transmembrane LRR protein DMA-1 promotes dendrite branching and growth in C. elegans. Nat Neurosci 15, 57-63.
Lodato, S., Molyneaux, B.J., Zuccaro, E., Goff, L.A., Chen, H.H., Yuan, W., Meleski, A., Takahashi, E., Mahony, S., Rinn, J.L., et al. (2014). Gene co-regulation by Fezf2 selects neurotransmitter identity and connectivity of corticospinal neurons. Nat Neurosci 17, 1046-54.
Martinat, C., Bacci, J.J., Leete, T., Kim, J., Vanti, W.B., Newman, A.H., Cha, J.H., Gether, U., Wang, H., and Abeliovich, A. (2006). Cooperative transcription activation by Nurr1 and Pitx3 induces embryonic stem cell maturation to the midbrain dopamine neuron phenotype. Proc Natl Acad Sci U S A 103, 2874-9.
Mello, C.C., Kramer, J.M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. EMBO J 10, 3959-70.
Mikeladze-Dvali, T., Wernet, M.F., Pistillo, D., Mazzoni, E.O., Teleman, A.A., Chen, Y.W., Cohen, S., and Desplan, C. (2005). The growth regulators warts/lats and melted interact in a bistable loop to specify opposite fates in Drosophila R8 photoreceptors. Cell 122, 775-87.
Mitani, S., Du, H., Hall, D.H., Driscoll, M., and Chalfie, M. (1993). Combinatorial control of touch receptor neuron expression in Caenorhabditis elegans. Development 119, 773-83.
Miyoshi, T., Maruhashi, M., Van De Putte, T., Kondoh, H., Huylebroeck, D., and Higashi, Y. (2006). Complementary expression pattern of Zfhx1 genes Sip1 and deltaEF1 in the mouse embryo and their genetic interaction revealed by compound mutants. Dev Dyn 235, 1941-52.
Nakatani, T., Minaki, Y., Kumai, M., and Ono, Y. (2007). Helt determines GABAergic over glutamatergic neuronal fate by repressing Ngn genes in the developing mesencephalon. Development 134, 2783-93.
Pereira, L., Kratsios, P., Serrano-Saiz, E., Sheftel, H., Mayo, A.E., Hall, D.H., White, J.G., LeBoeuf, B., Garcia, L.R., Alon, U., et al. (2015). A cellular and regulatory map of the cholinergic nervous system of C. elegans. Elife 4.
Poole, R.J., Bashllari, E., Cochella, L., Flowers, E.B., and Hobert, O. (2011). A Genome-Wide RNAi Screen for Factors Involved in Neuronal Specification in Caenorhabditis elegans. PLoS Genet 7, e1002109.
Russell, J., Vidal-Gadea, A.G., Makay, A., Lanam, C., and Pierce-Shimomura, J.T. (2014). Humidity sensation requires both mechanosensory and thermosensory pathways in Caenorhabditis elegans. Proc Natl Acad Sci U S A 111, 8269-74.
Sarin, S., O'Meara, M.M., Flowers, E.B., Antonio, C., Poole, R.J., Didiano, D., Johnston, R.J., Jr., Chang, S., Narula, S., and Hobert, O. (2007). Genetic screens for Caenorhabditis elegans mutants defective in left/right asymmetric neuronal fate specification. Genetics 176, 2109-30.
Sellin, J., Drechsler, M., Nguyen, H.T., and Paululat, A. (2009). Antagonistic function of Lmd and Zfh1 fine tunes cell fate decisions in the Twi and Tin positive mesoderm of Drosophila melanogaster. Dev Biol 326, 444-55.
Serrano-Saiz, E., Poole, R.J., Felton, T., Zhang, F., De La Cruz, E.D., and Hobert, O. (2013). Modular control of glutamatergic neuronal identity in C. elegans by distinct homeodomain proteins. Cell 155, 659-73.
Smith, C.J., O'Brien, T., Chatzigeorgiou, M., Spencer, W.C., Feingold-Link, E., Husson, S.J., Hori, S., Mitani, S., Gottschalk, A., Schafer, W.R., et al. (2013). Sensory neuron fates are distinguished by a transcriptional switch that regulates dendrite branch stabilization. Neuron 79, 266-80.
Sulston, J.E., and Horvitz, H.R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev Biol 56, 110-56.
Topalidou, I., and Chalfie, M. (2011). Shared gene expression in distinct neurons expressing common selector genes. Proc Natl Acad Sci U S A 108, 19258-63.
Topalidou, I., van Oudenaarden, A., and Chalfie, M. (2011). Caenorhabditis elegans aristaless/Arx gene alr-1 restricts variable gene expression. Proc Natl Acad Sci U S A 108, 4063-8.
Vandewalle, C., Van Roy, F., and Berx, G. (2009). The role of the ZEB family of transcription factors in development and disease. Cell Mol Life Sci 66, 773-87.

Wacker, I., Schwarz, V., Hedgecock, E.M., and Hutter, H. (2003). zag-1, a Zn-finger homeodomain transcription factor controlling neuronal differentiation and axon outgrowth in C. elegans. Development 130, 3795-805.

Way, J.C., and Chalfie, M. (1988). mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans. Cell 54, 5-16.

Weihe, U., Milan, M., and Cohen, S.M. (2001). Regulation of Apterous activity in Drosophila wing development. Development 128, 4615-22.

Wu, J., Duggan, A., and Chalfie, M. (2001). Inhibition of touch cell fate by egl-44 and egl-46 in C. elegans. Genes Dev 15, 789-802.

Xue, D., Finney, M., Ruvkun, G., and Chalfie, M. (1992). Regulation of the mec-3 gene by the C.elegans homeoproteins UNC-86 and MEC-3. EMBO J 11, 4969-79.

Xue, D., Tu, Y., and Chalfie, M. (1993). Cooperative interactions between the Caenorhabditis elegans homeoproteins UNC-86 and MEC-3. Science 261, 1324-8.

Zhang, Y., Ma, C., Delohery, T., Nasipak, B., Foat, B.C., Bounoutas, A., Bussemaker, H.J., Kim, S.K., and Chalfie, M. (2002). Identification of genes expressed in C. elegans touch receptor neurons. Nature 418, 331-5.

Zheng, C., Jin, F.Q., and Chalfie, M. (2015). Hox Proteins Act as Transcriptional Guarantors to Ensure Terminal Differentiation. Cell Rep 13, 1343-1352.

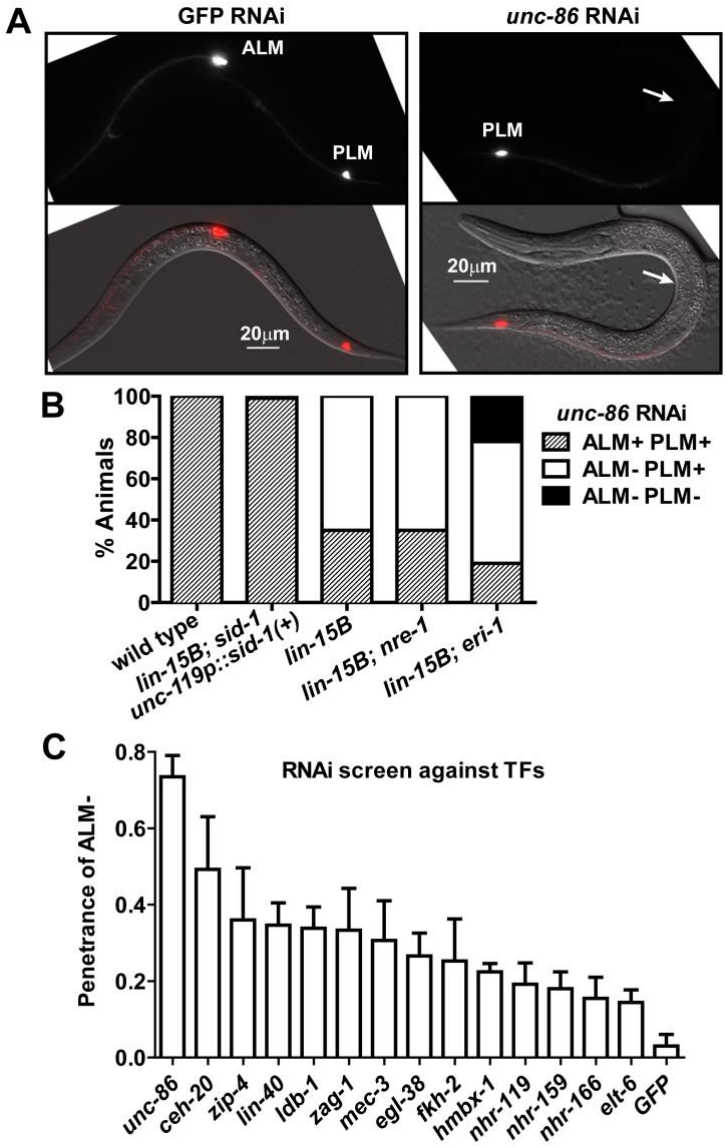
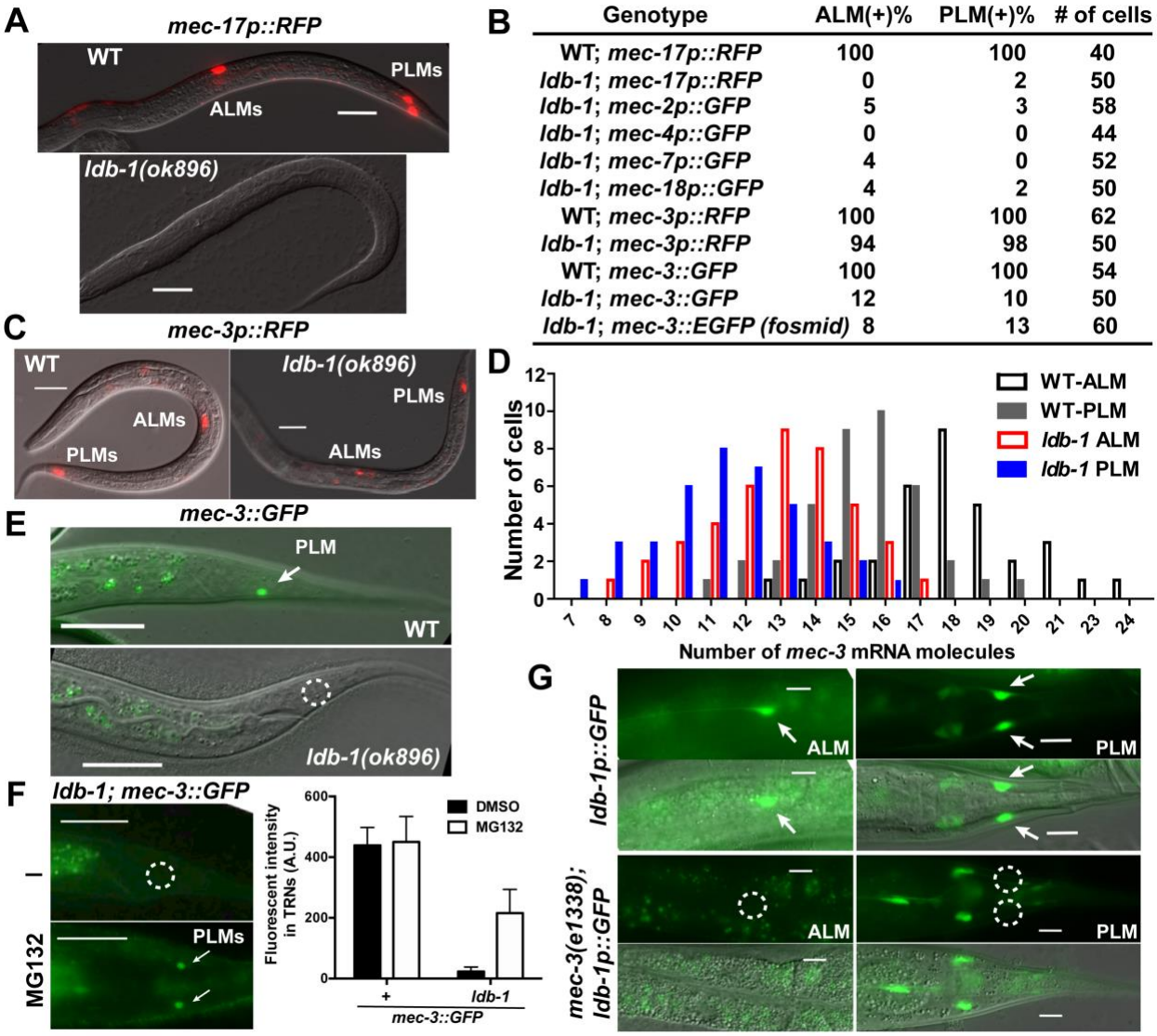
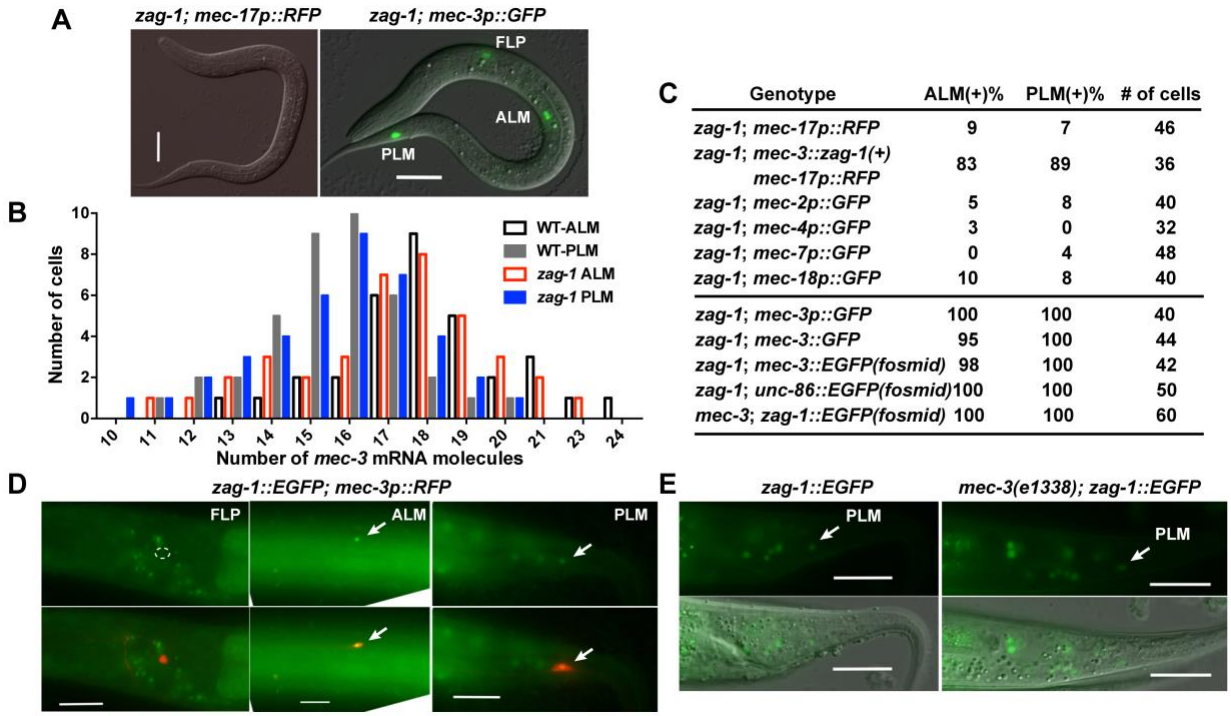


Figure 1. RNAi screen identifies positive regulators of TRN fate. (A) TU4429, *eri-1(mg366); lin-15B(n744); uIs134[mec-17p::RFP]* animals treated with RNAi against *unc-86* or GFP. (B) Percentage of animals that showed RFP or GFP expression in at least one ALM and one PLM (ALM+PLM+), in no ALM but at least one PLM (ALM-PLM+), and in no ALM and no PLM (ALM-PLM-) neurons, respectively. Strains tested for the efficiency of RNAi are TU4429, TU4396, *nre-1(hd20) lin-15B(hd126); uIs134[mec-17p::RFP]*, TU3595, *sid-1(pk3321) him-5(e1490); lin-15B(n744); uIs72[mec-18::GFP]*, and TU4301, *lin-15B(n744); uIs115[mec-17p::RFP]*. (C) The positive RNAi clones identified from the screen; $p > 0.05$ for all the positives using the data from all five rounds of screen.



Genotype	ALM(+)%	PLM(+)%	# of cells
WT; *mec-17p::RFP*	100	100	40
ldb-1; *mec-17p::RFP*	0	2	50
ldb-1; *mec-2p::GFP*	5	3	58
ldb-1; *mec-4p::GFP*	0	0	44
ldb-1; *mec-7p::GFP*	4	0	52
ldb-1; *mec-18p::GFP*	4	2	50
WT; *mec-3p::RFP*	100	100	62
ldb-1; *mec-3p::RFP*	94	98	50
WT; *mec-3::GFP*	100	100	54
ldb-1; *mec-3::GFP*	12	10	50
ldb-1; *mec-3::EGFP (fosmid)*	8	13	60

Figure 2. *ldb-1* is required for the expression of TRN fate. (A-C) The expression of TRN markers in wild-type and *ldb-1(ok896)* animals at the first larval (L1) stage. The penetrance for the expression of the TRN fate markers in ALM and PLM neurons are all 100% in the wild-type animals. (D) The number of *mec-3* transcripts in TRNs from wild-type and *ldb-1* animals from smFISH experiments. (E) The loss of the expression of *mec-3::GFP* (abbreviated for *mec-3p::mec-3::GFP::unc-54 3'UTR*) translational fusion in *ldb-1* animals. (F) The expression of a fosmid-based *mec-3* reporter *wgIs55[mec-3::EGFP]* in *ldb-1* mutants grown in the presence of 10 μM proteasome inhibitor MG132 in the NGM plates. (G) The loss of *ldb-1p::GFP* expression in undifferentiated TRN neurons from *mec-3* animals. Scale bars = 20 μm. Arrows point to cells expressing the marker, and dashed circles indicate the absence of expression; for penetrance, the same number of cells for each cell type were examined for a strain (the same applies to all other figures). Also see Figure S1.



Genotype	ALM(+)%	PLM(+)%	# of cells
zag-1; mec-17p::RFP	9	7	46
zag-1; mec-3::zag-1(+) mec-17p::RFP	83	89	36
zag-1; mec-2p::GFP	5	8	40
zag-1; mec-4p::GFP	3	0	32
zag-1; mec-7p::GFP	0	4	48
zag-1; mec-18p::GFP	10	8	40
zag-1; mec-3p::GFP	100	100	40
zag-1; mec-3::GFP	95	100	44
zag-1; mec-3::EGFP(fosmid)	98	100	42
zag-1; unc-86::EGFP(fosmid)	100	100	50
mec-3; zag-1::EGFP(fosmid)	100	100	60

Figure 3. *zag-1* is required for the expression of TRN markers independently of *mec-3*. (A) The expression of *mec-17p::RFP* and *mec-3p::GFP* reporters in *zag-1(hd16)* animals at L1 stage. (B) The number of *mec-3* transcripts in TRNs from wild-type and *zag-1* animals from smFISH experiments. (C) The penetrance for the expression of the TRN fate markers, and *mec-3*, *unc-86*, and *zag-1* reporters in ALM and PLM cells of *zag-1(hd16)* animals at L1 stage. Those reporters are expressed in 100% ALM and PLM cells in the wild-type animals. (D) The expression of fosmid-based reporter *zag-1::EGFP* in TRNs but not FLPs. (E) *zag-1:EGFP* expression in *mec-3* mutants. Scale bars = 20 μm. Also see Figure S2.

A

Genotype	FLP(+)%	ALM(+)%	PLM(+)%	# of cells
WT; *mec-17p::GFP*	0	100	100	50
egl-44(n1080); mec-17p::GFP	84	100	100	58
egl-46(gk692); mec-17p::GFP	75	100	100	56
egl-46(n1075); mec-17p::GFP	30	100	100	60
zag-1; egl-44; mec-17p::GFP	80	90	88	40
zag-1; egl-46; mec-17p::GFP	69	84	89	36
egl-44::GFP	100	0	0	40
egl-46::GFP	100	0	0	44
egl-44; egl-46::GFP	5	0	0	42
zag-1; egl-44::GFP	100	81	78	32
zag-1; egl-46::GFP	100	76	82	38

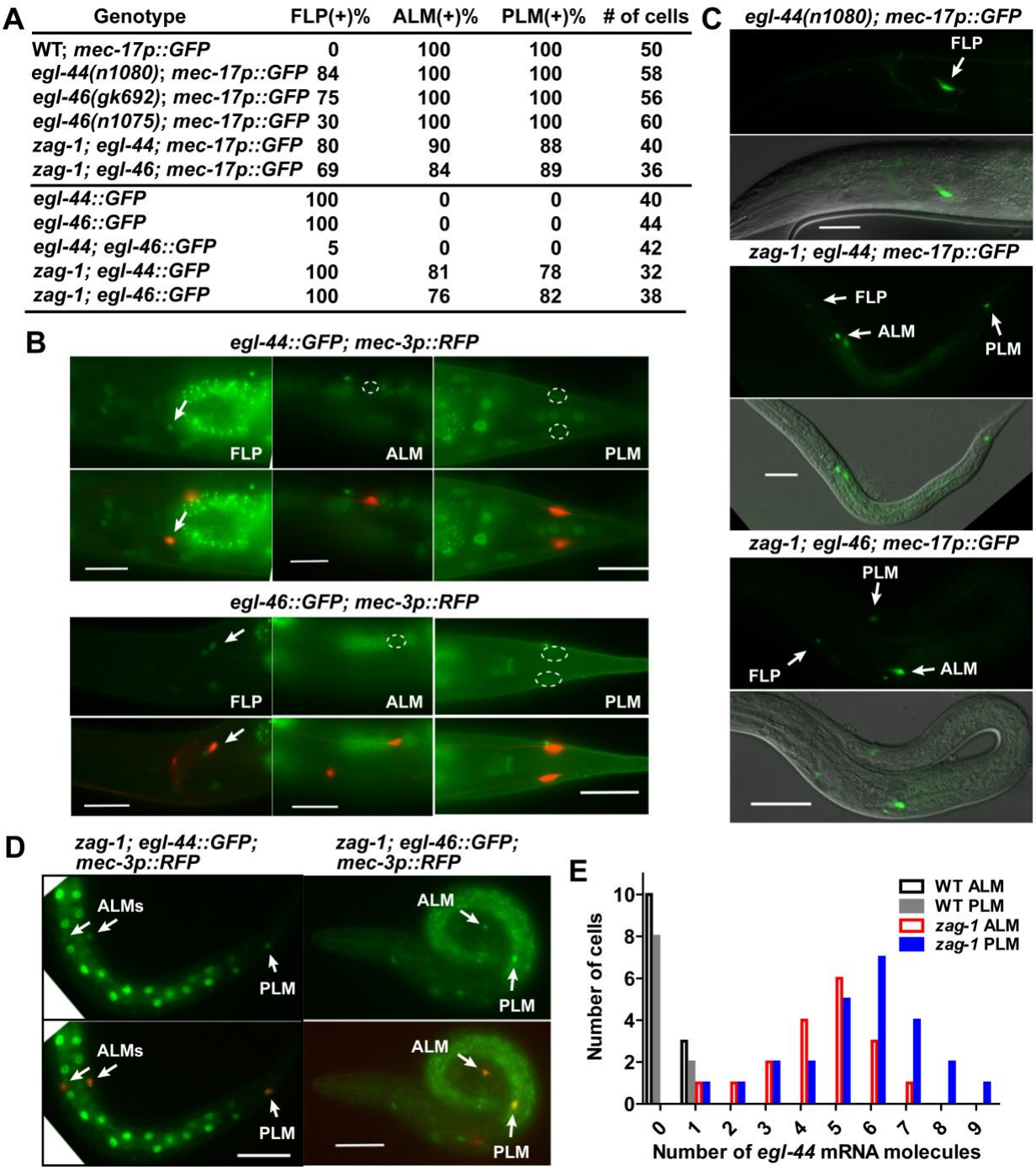
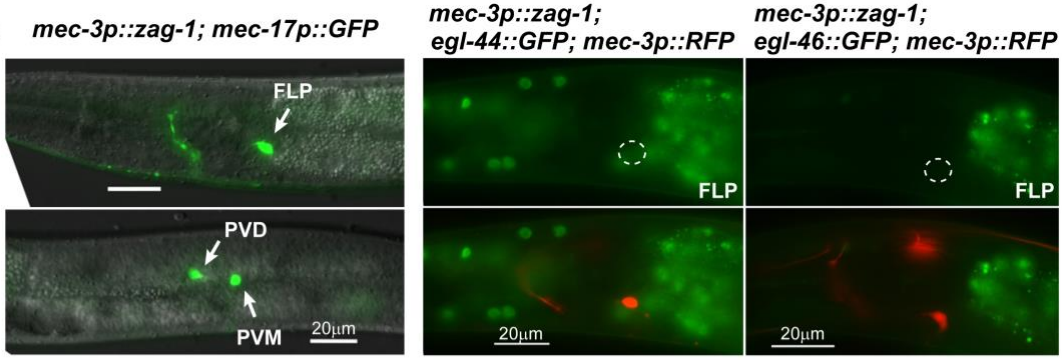


Figure 4. ZAG-1 promotes TRN fate by repressing *egl-44* and *egl-46*. (A) The expression of TRN marker *mec-17p::GFP* in FLP neurons of *egl-44* animals and in TRNs in *egl-44; zag-1* and *zag-1; egl-46* animals. (B) Penetrance of the expression of various reporters. (C-D) The expression of *egl-44::GFP* and *egl-46::GFP* reporters in FLPs but not TRNs of wild-type animals and in TRNs of *zag-1* animals. (E) The number of *egl-44* mRNA molecules in TRNs of wild-type and *zag-1* animals. Also see Figure S3.

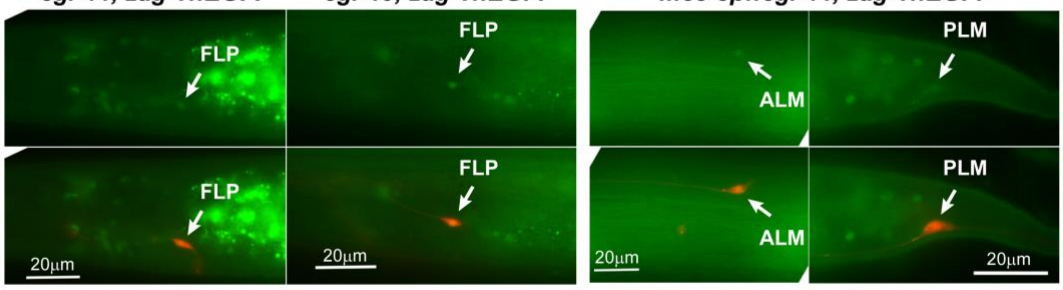
A



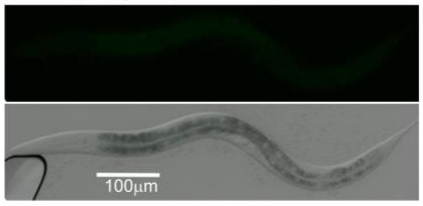
B

Genotype	FLP(+)%	ALM(+)%	PLM(+)%	# of cells	
mec-3p::zag-1; mec-17p::GFP	90	100	100	50	
mec-3p::ldb-1; mec-17p::GFP	0	100	100	50	
mec-3p::zag-1; egl-44::GFP	2	0	0	60	
mec-3p::zag-1; egl-46::GFP	0	0	0	52	
egl-44; zag-1::EGFP	82	100	100	60	
egl-46; zag-1::EGFP	75	100	100	56	
mec-3p::egl-44; mec-17p::GFP	0	0	0	50	
mec-3p::egl-46; mec-17p::GFP	0	100	100	54	
mec-3p::egl-44; egl-46::GFP	100	93	95	58	
mec-3p::egl-44; zag-1::EGFP	0	88	92	60	
	0	55	61	60	(%weak exp.)
mec-3p::egl-44; mec-3p::zag-1; mec-17p::GFP	87	90	96	68	
mec-3p::egl-44; mec-3p::zag-1 mec-3p::egl-46; mec-17p::GFP	2	0	0	62	
mec-3p::egl-44; mec-3p::zag-1; egl-46::GFP	3	3	0	30	
egl-46; mec-3p::egl-44; mec-17p::GFP	44	60	67	48	

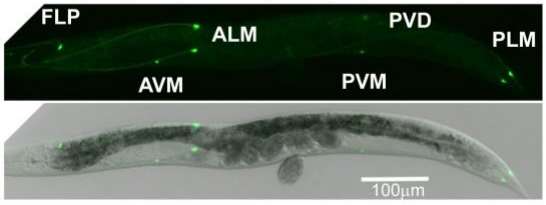
C



D



E



F

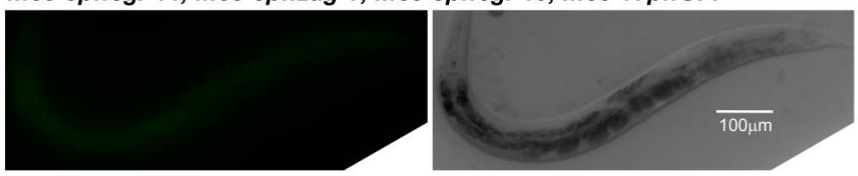
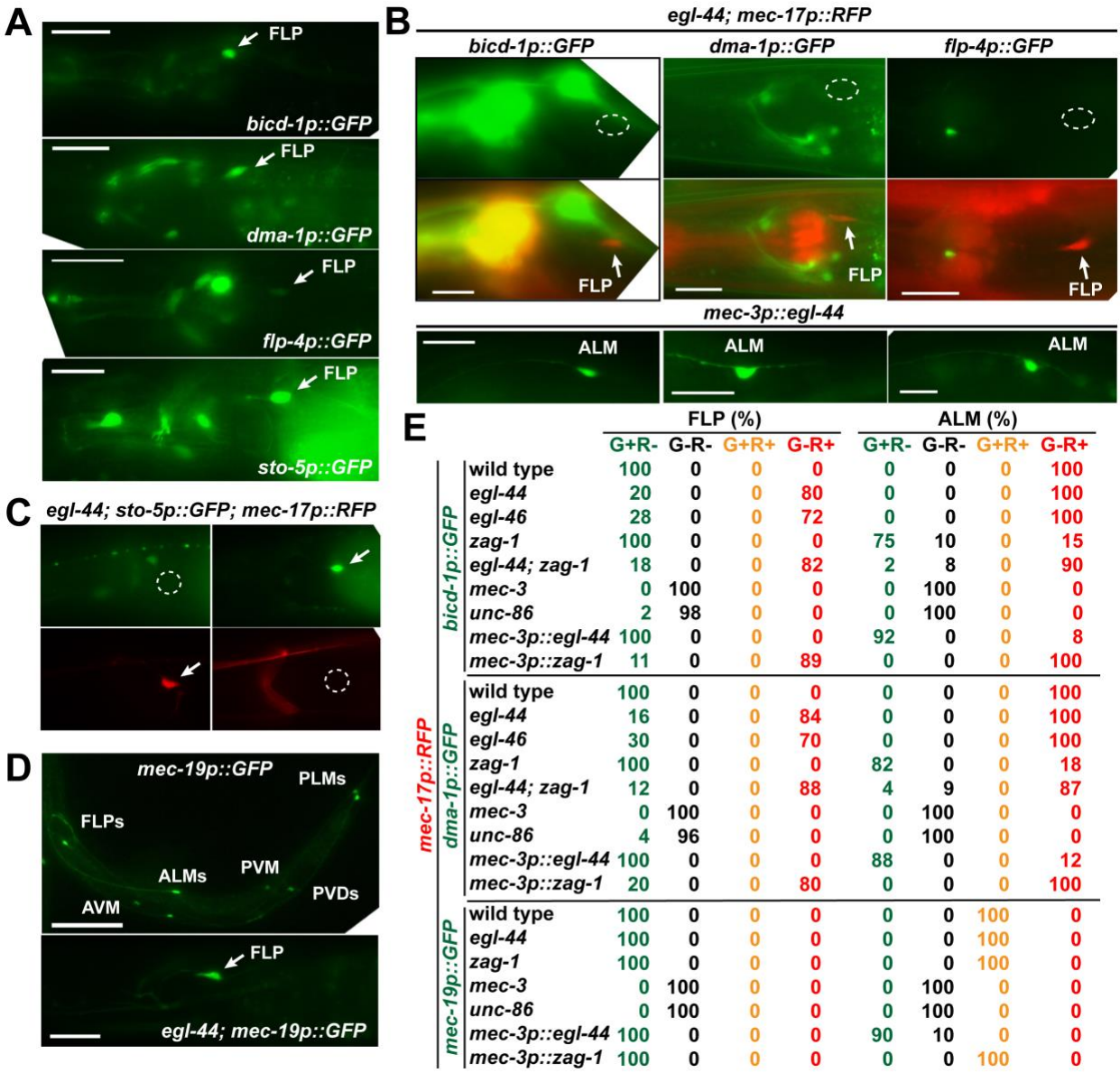


Figure 5. Mutual inhibition between *zag-1* and *egl-44* regulate TRN fate decision. (A) The activation of TRN marker and the loss of the expression of *egl-44* and *egl-46* reporters in FLP neurons of animals carrying *mec-3p::zag-1* transgene. (B) The percentage of cells expressing the indicated markers in various strains. (C) The expression of *zag-1::EGFP* in FLP neurons of *egl-44* and *egl-46* mutants, and weak expression of *zag-1* reporter in animals misexpressing *egl-44* from *mec-3* promoter. (D-F) The expression of TRN marker *mec-17p::GFP* in animals carrying various transgenes. Also see Figure S4.



E

		FLP (%)				ALM (%)			
		G+R-	G-R-	G+R+	G-R+	G+R-	G-R-	G+R+	G-R+
bicd-1p::GFP	wild type	100	0	0	0	0	0	0	100
	egl-44	20	0	0	80	0	0	0	100
	egl-46	28	0	0	72	0	0	0	100
	zag-1	100	0	0	0	75	10	0	15
	egl-44; zag-1	18	0	0	82	2	8	0	90
	mec-3	0	100	0	0	0	100	0	0
	unc-86	2	98	0	0	0	100	0	0
	mec-3p::egl-44	100	0	0	0	92	0	0	8
	mec-3p::zag-1	11	0	0	89	0	0	0	100
mec-17p::RFP; dma-1p::GFP	wild type	100	0	0	0	0	0	0	100
	egl-44	16	0	0	84	0	0	0	100
	egl-46	30	0	0	70	0	0	0	100
	zag-1	100	0	0	0	82	0	0	18
	egl-44; zag-1	12	0	0	88	4	9	0	87
	mec-3	0	100	0	0	0	100	0	0
	unc-86	4	96	0	0	0	100	0	0
	mec-3p::egl-44	100	0	0	0	88	0	0	12
	mec-3p::zag-1	20	0	0	80	0	0	0	100
mec-19p::GFP	wild type	100	0	0	0	0	0	100	0
	egl-44	100	0	0	0	0	0	100	0
	zag-1	100	0	0	0	0	0	100	0
	mec-3	0	100	0	0	0	100	0	0
	unc-86	0	100	0	0	0	100	0	0
	mec-3p::egl-44	100	0	0	0	90	10	0	0
	mec-3p::zag-1	100	0	0	0	0	0	100	0

Figure 6. EGL-44/EGL-46 simultaneously induces FLP genes and suppresses TRN genes. (A-B) Various FLP fate reporters were expressed in wild-type FLP neurons but not in *egl-44* mutants. (C) The expression of FLP and TRN fate markers were mutually exclusive in FLP neurons of *egl-44* mutants. (D) The expression of mec-19 was not affected by *egl-44* mutation. (E) Percentage of FLP and ALM neurons expressing the green (G) and red (R) markers in various strains. The number of cells > 40. Also see Figure S5.

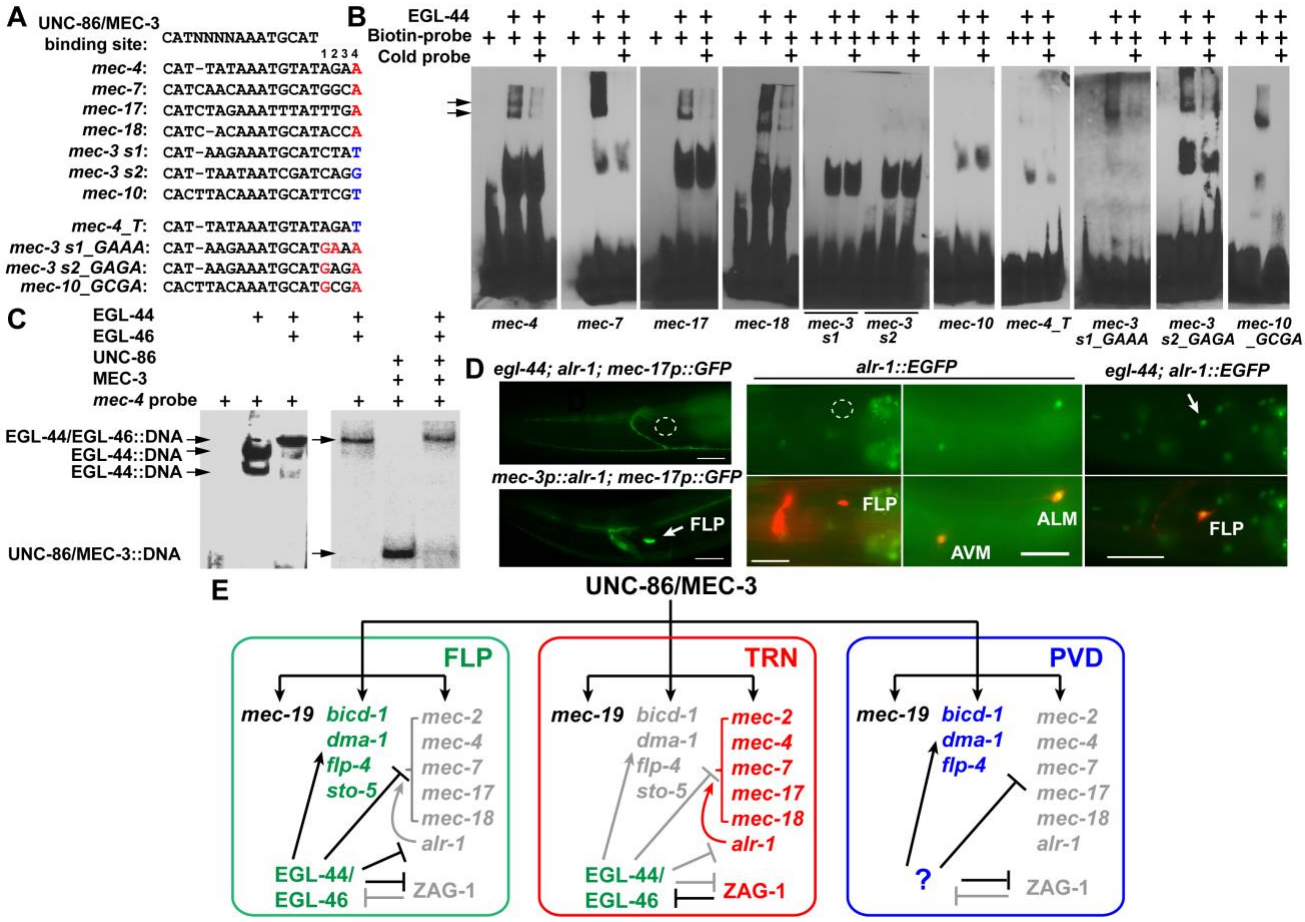


Figure 7. EGL-44/EGL-46 inhibits the expression of TRN genes by outcompeting UNC-86/MEC-3 for the binding of the same *cis*-regulatory elements. (A) The alignment of various *cis*-regulatory motifs tested in electrophoretic mobility shift (EMSA) assays. Positions 1-4 were assigned to the four nucleotides following the consensus UNC-86/MEC-3 binding site (Zhang et al., 2002). Nucleotides in red were considered important for EGL-44 binding and nucleotides in blue were responsible for the lack of EGL-44 binding. (B) The binding of recombinant EGL-44 proteins to various probes in EMSA assays. Arrows point to the band of EGL-44::DNA complexes. (C) The binding of EGL-44/EGL-46 with *mec-4* probes in the presence or absence of UNC-86/MEC-3. (D) The expression of TRN fate marker in *egl-44; alr-1* double mutants and animals misexpressing ALR-1 from the *mec-3* promoter and the expression of a fosmid-based *alr-1* reporter *wgIs200[alr-1::EGFP]* in wild-type animals and *egl-44* mutants. (E) A model for the regulatory mechanisms controlling the cell fate specification among FLP, PVD, and TRN fates. Gene names in green, red, and blue indicate genes expressed in FLP, TRN, and PVD neurons, respectively; black names for genes commonly expressed in all three types and grey name for genes repressed. Also see Figure S6 and S7.